

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Ecologie et Biodiversité

École doctorale GAIA

Unité de recherche AGAP

En partenariat international avec l'Université Libanaise, LIBAN

Titre de la thèse

**Structure et dynamique de la diversité génétique de l'amandier
cultivé au Liban : facteurs biologiques et anthropiques**

Présentée par Bariaa HAMADEH

Le 28 Septembre 2018

Sous la direction de Hélène JOLY
et Lamis CHALAK

Devant le jury composé de

Bruno FADY, directeur de recherche, INRA

Thierry ROBERT, maître de conférences, Université Paris-Sud XI

Daniel PRAT, professeur, Université de Lyon

Romain SIMENEL, chargé de recherche, IRD

Hélène JOLY, chercheur, CIRAD

Lamis CHALAK, professeur, Université Libanaise

Pascale MAIZI, maître de conférences, SupAgro

Examineur, Président du jury

Rapporteur

Rapporteur

Examineur

Directrice de thèse

Directrice de thèse

Co-encadrante de thèse, Membre invité



UNIVERSITÉ
DE MONTPELLIER

| | |
|--|------------|
| INTRODUCTION GENERALE | 4 |
| I. Contexte et enjeu | 5 |
| II. Effet de l'homme sur la dynamique de la diversité des plantes cultivées | 7 |
| II.1 la domestication | 7 |
| II.2 Interaction entre homme et forces évolutives | 10 |
| II.2.1 Sélection | 10 |
| II.2.2 Migration/diffusion | 12 |
| II.2.3 Dérive | 13 |
| II.3 Facteurs culturels et diversité des plantes cultivées | 14 |
| III. Problématique de la thèse | 15 |
| III.1 Modèle d'étude: cadre bio géographique de la thèse | 15 |
| III.1.1 L'amandier | 15 |
| III.1.2 Le Liban | 19 |
| III.2 Questions de recherche et démarche | 21 |
| III.3 Plan de la thèse | 22 |
| CHAPITRE 1 | 28 |
| Impact of ancient and modern human-mediated gene flow on the genetic structure of cultivated almonds in the Mediterranean Basin | 29 |
| CHAPITRE 2 | 66 |
| Taxonomie locale et caractérisation de la diversité de l'amandier cultivé au Liban | 67 |
| CHAPITRE 3 | 100 |

| | |
|--|------------|
| Evolution of almond genetic diversity and farmer practices in Lebanon: Impacts of the diffusion of a graft-propagated cultivar in a traditional system based on seed-propagation----- | 101 |
|--|------------|

| | |
|---|------------|
| DISCUSSION GENERALE ET PERSPECTIVES----- | 142 |
|---|------------|

| | |
|---|------------|
| Dynamique de la diversité génétique et diffusion des amandiers cultivés au Liban, influence des savoirs et des pratiques: regards croisés génétique – anthropologie----- | 143 |
|---|------------|

| | |
|--|------------|
| 1) Dynamique d'introduction, diversité nommée et diversité variétale----- | 144 |
|--|------------|

| | |
|--|------------|
| 2) Dynamique de la diversité génétique, diffusion et pratiques de propagation---- | 147 |
|--|------------|

| | |
|---|------------|
| 3) Quelques implications pour la conservation----- | 151 |
|---|------------|

| | |
|---|------------|
| <i>i) Conservation ex situ-----</i> | <i>152</i> |
|---|------------|

| | |
|---|------------|
| <i>ii) Conservation in situ-----</i> | <i>152</i> |
|---|------------|

| | |
|--|------------|
| <i>iii) Conservation des ressources génétiques spontanées-----</i> | <i>153</i> |
|--|------------|

| | |
|--|------------|
| <i>iv) Utilisation durable-----</i> | <i>153</i> |
|--|------------|

| | |
|---------------------|------------|
| ANNEXES----- | 155 |
|---------------------|------------|

Remerciements

Mes premiers remerciements s'adressent tout d'abord à mes encadrantes, Lamis Chalak et Hélène Joly, pour leur disponibilité, et leurs conseils, tout au long de cette thèse. Merci pour votre soutien et de m'avoir épaulé à différentes périodes de la thèse. Je suis extrêmement sensible à vos qualités humaines d'écoute et de compréhension. Hélène, je te remercie encore, entre les lignes.

Je remercie aussi ma co-encadrante Pascale Maizi pour ses maints conseils pour les aspects anthropologiques. Je souhaite exprimer ma gratitude à Geo Coppens pour les discussions enrichissantes et sa contribution. Je vous remercie pour votre accueil chaleureux à chaque fois que j'ai sollicité votre aide.

Je tiens à remercier les membres du jury d'avoir accepté d'évaluer mon travail : Thierry Robert, Daniel Prat, Hélène Rizk, Bruno Fady et Romain Simenel.

Je remercie également Bouchaib Kadari, Jean-Frédéric Terral et Xavier Perrier, pour les discussions constructives et d'avoir guidé mes réflexions lors de mes comités de suivi de thèse. Eric Garine pour les discussions et la bibliographie autour de la nomenclature et taxinomie vernaculaire.

Je tiens à remercier, Michel Afram, Directeur de l'IRAL, qui m'a soutenu et m'a permis de bénéficier des facilités de l'IRAL ainsi que d'un support financier et logistique jusqu'au juin 2016. Je tiens aussi à remercier Philippe Jarne et Yildiz Aumeeruddy-Thomas qui m'ont accueillie au CEFÉ et Claire Billot qui m'a accueillie dans l'équipe DDSE de l'UMR AGAP pour la poursuite de mon travail.

Je tiens à remercier Fawaz el Omar (Doyen de l'EDST, Université Libanaise) et Marc Bouvy (directeur de l'ED GAIA, Université de Montpellier) ainsi que Zeinab Ibrahim (EDST, UL) et Véronique Dubois (service Handiversité, UM), d'avoir facilité ma reprise de thèse.

Merci surtout aux agriculteurs et pépiniériste libanais qui ont participé à cette étude et qui ont partagé une partie de leurs connaissances !

Je remercie tous ceux qui m'ont soutenu à différentes périodes de la thèse. Au CEFÉ, Richard Joffre, Finn Kjellberg, Edmon Dounias, Sophie Caillon, Thierry Winkel, et bien d'autres personnes ! Laure merci de m'avoir guidé au laboratoire. Au CIRAD, l'équipe DDSE et surtout Luc Baudouin, Aziz Saidou et Mathieu Thomas de m'avoir aidé à la compréhension de certains logiciels. Guylaine, merci surtout pour ton coup de main de dernière minute.

Merci à mes amis Sali, Oula, Suzy, Clélia, Lucie, Aziz, Menad et bien d'autres, pour ce que vous êtes. Jack, merci pour les papillons ! Mohammad merci pour m'avoir accompagné sur le terrain sur l'amandier en 2014!

Danièle, Jean-François, Elsa ma très belle famille, merci pour vos encouragements et votre soutien affectif. Je remercie aussi Joelle pour le temps libéré.

Maman et Papa, le merci ne suffit pas car cette thèse est le fruit de votre soutien constant. Mounah, Athar et Mohammad, grand merci pour votre soutien affectif.

Enfin, Jules et Léon Yahia, je vous remercie d'avoir rempli mon quotidien de bonheur. Je vous remercie vous deux pour votre soutien et d'avoir été à mes côtés dans cette traversée...

Préambule

Cette thèse s'inscrit sous la thématique générale de l'évaluation de la diversité génétique des arbres fruitiers dans le cadre de la Stratégie Nationale du Liban pour la Conservation et la Gestion des Ressources Phytogénétiques pour l'Alimentation et l'Agriculture. De plus, cette thèse est une contribution à l'observatoire Libano-Français de l'environnement O-life. Les travaux de la thèse ont démarré dans le cadre du projet FruitMed 901-007 soutenu par la fondation Agropolis et du projet Franco-Libanais CEDRE réf. 09 E F6/L 24 mis en œuvre en France par le ministère de l'Europe et des Affaires étrangères (MEAE) et le Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation (MESRI), et au Liban, par le Ministère de l'Éducation et de l'Enseignement Supérieur. Ma thèse a été initialement lancée pendant que j'occupais le poste de chercheur à l'Institut National de Recherches Agronomiques du Liban (IRAL), et où j'ai bénéficié d'un salaire régulier et de facilités diverses (Novembre 2012 - Juin 2016). Pendant la thèse j'ai aussi bénéficié d'un soutien pour les jeunes chercheurs du programme BioDivMex (Septembre 2018).

Il s'agit d'une thèse en cotutelle entre l'Université Libanaise et l'Université de Montpellier. Le travail s'est déroulée en alternance entre Montpellier et le Liban dans différents laboratoires d'accueil : de novembre 2012 jusqu'à décembre 2014 dans l'UMR CEFE (Centre d'Ecologie Fonctionnelle et Evolutive), équipe Interactions Bio-Culturelles, puis de février 2015 à sa soutenance dans l'UMR AGAP (Amélioration Génétique et Adaptation des Plantes), équipe Dynamique de la Diversité, Sociétés et Environnement. La collecte du matériel végétal s'est déroulée au Liban dans le cadre de mes fonctions à l'IRAL, en collaboration avec l'Université libanaise. Le marquage moléculaire a été réalisé au laboratoire de biologie moléculaire de l'UMR CEFE et dans celui de la plateforme technique du Labex CeMed de l'Université de Montpellier. L'analyse des données, tant génétiques qu'anthropologiques, et l'interprétation des résultats ont été effectuées au CEFE puis à AGAP. La rédaction, des articles comme du document de thèse, s'est faite à l'UMR AGAP, Campus CIRAD Lavalette.

"La vie est d'abord une affaire de diversité"

Robert Barbault
Un éléphant dans un jeu de quille
Seuil, 2006

Introduction générale

I. Contexte et enjeu

Chaque espèce contient une diversité représentée par des différences entre les individus au niveau des gènes, de la taille, des formes, des couleurs, des rythmes de croissance, ainsi que des adaptations diverses à l'environnement et à ses changements. En 1992, la convention sur la diversité biologique (CBD) a mis en exergue la conservation et l'utilisation de la diversité biologique ; cette convention met l'accent sur le maintien et l'évolution des espèces, dans le temps et dans l'espace.

Depuis 12000 ans, l'homme en passant du chasseur-cueilleur à l'agriculteur a appris à utiliser les plantes de son environnement et a modelé la diversité des espèces domestiquées. Le rôle des agriculteurs dans la création et la préservation de la diversité des ressources génétiques cultivées a été reconnu officiellement par la FAO en 1996: "Plant genetic resources for food and agriculture are the product of natural evolution and human intervention. We acknowledge the roles played by generations of men and women farmers and plant breeders, and by indigenous and local communities, in conserving and improving plant genetic resources". Cette reconnaissance souligne donc que la diversité actuelle est le produit d'un processus d'évolution et de diversification des espèces, dont l'Homme impose une grande part des règles. La diversité des espèces cultivées résulte d'une série d'évènements de domestication, de flux de gènes entre compartiments sauvages et cultivés, d'effets de sélection adaptative naturelle mais aussi de sélection et de dispersions géographiques à grandes échelles par l'Homme, souvent sur de longues périodes. Les processus impliqués dans la structuration de la diversité des espèces cultivées sont variables non seulement en fonction de la biologie de l'espèce étudiée mais aussi de la perception et des décisions des agriculteurs car la dynamique de la diversité de ces espèces est fortement liée à leurs savoirs et savoir-faire.

La compréhension de ces dynamiques d'évolution nécessite une approche interdisciplinaire entre la biologie et les sciences sociales, prenant en compte les mécanismes liés aux sociétés humaines comme acteurs, passés ou contemporains. La compréhension des différents effets des facteurs humains depuis la domestication, la diffusion jusqu'à la sélection actuelle, sur la structure de la diversité génétique est donc essentielle pour évaluer la dynamique évolutive des espèces cultivées et leur diversification.

Selon les contextes agricoles la diversité des formes cultivées répond à une diversité de critères d'identification et de caractérisation, et plus généralement à une diversité des demandes sociales. La diversité des espèces cultivées est importante pour la sécurité alimentaire et l'équilibre nutritionnel. Les capacités d'adaptation des cultivars traditionnels deviennent encore plus nécessaires face aux contraintes liées au changement climatique. Ainsi, ils contribuent à des systèmes de production plus résilients et réduisent le risque de perte des récoltes.

L'évaluation de la diversité des espèces cultivées est primordiale afin de définir les priorités et les méthodes à adopter pour mettre en place un programme de conservation de leur diversité (Stockwell, Hendry and Kinnison, 2003). Les études récentes sur la diversité des espèces cultivées prennent en compte la richesse socioculturelle qui relie la nature à l'Homme, responsable décisif de la diversité des espèces cultivées. McKey et al. (2012), dans une revue de littérature sur les études des systèmes traditionnels de culture du manioc (*Manihot esculenta*) amérindien, ont montré que les agriculteurs associent les modes de propagation clonale et sexuée. Ces travaux montrent que les pratiques des agriculteurs permettent de conserver le potentiel adaptatif de l'espèce. Ces agriculteurs se basent sur leurs connaissances et leurs réseaux sociaux dans l'installation et la conduite de leurs cultures. Des études effectuées sur d'autres espèces telles que le sorgho (*Sorghum bicolor*) (Barnaud et al., 2007;

Labeyrie et al., 2016) et l'olivier (*Olea europaea*) (Haouane, 2012) ont lié des approches anthropologiques aux analyses génétiques pour une meilleure compréhension de l'influence des pratiques des agriculteurs sur la diversité génétique. Les activités sociales contribuent à la structuration génétique des populations de plantes domestiquées et nous imposent de considérer la diversité sociale comme une dimension structurante des interactions génotype-environnement (Leclerc and Coppens, 2012). L'objectif principal de cette thèse est de comprendre les processus de diversification de l'amandier (*Prunus dulcis*) au Liban et notamment comment les facteurs anthropiques et biologiques interagissent et façonnent sa diversité génétique.

II. Effet de l'homme sur la dynamique de la diversité des plantes cultivées

La diversité des plantes cultivées résulte de l'interaction entre les facteurs biologiques, environnementaux, historiques et sociaux. Les sociétés humaines ont agi directement sur la sélection et la migration des plantes cultivées.

II.1 la domestication

La domestication est un processus évolutif opérant sous l'influence des activités humaines (Harlan, 1992). Les changements évolutifs peuvent diminuer la valeur adaptative d'une plante dans la nature mais l'augmenter dans les environnements anthropisés (Harlan, 1992 ; Zeder et al., 2006). La manipulation des plantes par les communautés néolithiques a impliqué une nouvelle perception de la terre et des plantes, et un changement dans la relation entre culture et nature ; d'après Zeder et al. (2006), la domestication est la forme unique de mutualisme qui s'est développé entre les sociétés humaines et les populations de plantes.

En termes biologiques, la domestication se réfère aux changements génétiques qui sous-tendent les principaux changements de phénotypes; un ensemble de traits phénotypiques

appelé «syndrome de domestication», caractérise les plantes sélectionnées par l'homme, il est absent des parents sauvages. Ces changements sont liés à des traits morphologiques et physiologiques d'intérêt concernant l'usage (par exemple, augmentation de la taille des graines chez les céréales et du fruit chez la tomate *Solanum lycopersicum*), l'adaptation à la culture (par exemple, rachis non déhiscent ou absence de dormance) (Glémin and Bataillon, 2009). Le syndrome de domestication peut évoluer sur plusieurs générations tant que l'espèce est soumise à des pressions de sélection pour fixer le trait recherché (Fuller, 2007). Il peut aussi évoluer pendant une période assez courte comme pour le kiwi (*Actinidia deliciosa*) qui a été domestiqué il y a seulement une centaine d'année (Meyer, DuVal, and Jensen, 2012). En termes culturels, la domestication est un épisode basé sur une décision et une action de suivi par laquelle les humains ont choisi certaines espèces et ont sélectionné des plantes à caractères désirés qu'ils voulaient cultiver (Abbo et al., 2013).

La domestication des plantes s'est déroulée indépendamment une dizaine de fois dans différentes régions du monde ; dès 10 000 ans avant J.C. au Proche-Orient et jusqu'à 2 000 ans avant J.C. en Afrique de l'Ouest (Diamond, 2002 ; Purugganan and Fuller, 2009). La domestication des arbres a généralement été initiée plus tardivement et a eu un impact plus faible sur la structure du génome car relativement peu de générations se sont écoulées jusqu'à aujourd'hui (Miller and Gross, 2011). Les théories visant à expliquer les origines et le développement de l'agriculture ont pris en compte des facteurs allant de changements climatiques et démographiques aux pratiques culturelles et aux croyances religieuses (Harlan, 1992). Des études archéobotaniques suggèrent la présence d'une période de pré-domestication ; par exemple, la culture de céréales avant la domestication aurait duré pendant des siècles dans le croissant fertile (Weiss, Kislev and Hartmann, 2006 ; Willcox, Fornite and Herveux, 2007). Des études plus récentes ont montré que le processus de domestication s'est effectué sur une longue période, par exemple, l'acquisition du rachis non

déhiscent et l'augmentation de la taille des graines chez les céréales se sont graduellement effectuées sur une période de 2000 à 4000 ans (Tanno and Willcox, 2012). Zohary and Spiegel-Roy (1975) suggèrent que la domestication est plus consciente pour les arbres fruitiers (olivier, pommier, palmier dattier) et la vigne que pour les plantes annuelles, car les traits sélectionnés ont été maintenu par propagation clonale en général. Pour certains arbres fruitiers comme l'amandier, la domestication a été réalisée par propagation sexuée.

La diversification implique la diffusion et l'adaptation des espèces domestiquées à différentes conditions éco-géographiques et culturelles (Meyer and Purugganan, 2013). Cette étape a conduit à une divergence phénotypique et génétique entre des populations domestiquées d'une même espèce. Selon Meyer and Purugganan (2013), la diversification peut être considérée comme une série de plusieurs étapes de pressions sélectives variées : i) la multiplication *in situ* de populations porteuses de traits préalablement sélectionnés, ii) l'adaptation des populations de l'espèce domestiquée à différentes conditions environnementales et culturelles à la suite de la diffusion, et iii) l'amélioration pour avoir une meilleure production qualitative et quantitative. Pour certaines plantes domestiquées la dernière étape correspond à l'amélioration variétale menée depuis deux siècles par des sélectionneurs spécialisés et non plus seulement par les agriculteurs eux-mêmes. Bien que la domestication ait toujours été un processus dynamique, les innovations des deux siècles derniers entraînent des modifications drastiques du processus de diversification pour certaines espèces. Ces espèces sont soumises à de nouvelles pressions liées à la mise en place de programmes de sélection orientés vers de nouvelles utilisations des cultures existantes ou au déplacement de ces cultures vers de nouveaux environnements (Meyer and Purugganan, 2013).

La structure actuelle de la diversité génétique d'une plante cultivée est le résultat de son histoire évolutive ; celle-ci dépend directement de la biologie de cette espèce et de son interaction avec l'Homme. Miller et Gross (2011) soulignent que la diversité génétique des plantes cultivées annuelles est limitée par rapport à celle des plantes pérennes au niveau spécifique. Ceci peut s'expliquer d'une part, car la domestication des pérennes est plus récente et d'autre part, par la durée du cycle de vie qui limite le nombre de génération entre la phase de domestication et les observations actuelles ; cela est confirmé pour l'olivier (Haouane, 2012), pour le pommier (Cornille et al., 2012), pour le cerisier (Mariette et al., 2010) et pour l'amandier (Delplancke et al., 2013).

II.2 Interaction entre homme et forces évolutives

L'Homme tient une place centrale dans la structure et la dynamique de la diversité génétique des espèces domestiquées. Les plantes sont soumises aux forces évolutives qui sont influencées à la fois par les forces évolutives naturelles et par les migrations, sélections et dérives qu'entraîne l'agriculteur lors de l'installation de son verger. L'intensité de l'impact évolutif des pratiques agricoles dépend de leurs interactions avec la biologie de l'espèce cultivée.

II.2.1 Sélection

«Usually good variety of a native plant might attract the attention of some wise old savage; and he would transplant it, or sow its seeds. ... As the soil near the novels of the natives would often be in some degree manure, improved varieties would sooner or later arise. »
Darwin (1868), p.327

La sélection naturelle est la compétition pour la survie et la reproduction entre populations qui résulte des variations génétiques. Un individu ou une population doit avoir des caractéristiques qui lui permettent de survivre et de se reproduire pour se maintenir. Chez les plantes cultivées, les agriculteurs sélectionnent des individus en se basant sur des critères qui les intéressent. Cette sélection interagit avec les pressions de sélection naturelles.

Selon Darwin (1859) dans «The origin of species», la domestication s'est basée sur deux types de sélection, la sélection consciente et la sélection inconsciente. Les agriculteurs ont sélectionné consciemment des phénotypes d'intérêt, tandis qu'ils ont négligés ou activement éliminés les phénotypes moins souhaitables jusqu'à ce que leur fréquence diminue dans la population. De plus, et de manière généralement inconsciente, l'Homme a modifié les pressions de sélection en déplaçant les plantes de la nature vers des environnements plus contrôlés. Ceci a conduit à l'émergence de phénotypes qui ont une faible survie dans l'environnement naturel (Zohary, 2004). Cette sélection empirique, agissant de façon continue, a conduit à une variabilité phénotypique plus importante chez le taxon cultivé que chez le taxon sauvage. En effet, l'interaction de la sélection naturelle, de la sélection humaine pendant des temps longs ont conduit à la création de nombreux cultivars.

Depuis deux siècles, un nouveau type de sélection « moderne » faite par des scientifiques/sélectionneurs s'est développé. Des programmes d'amélioration variétale sont mis en place par des généticiens qui ont effectués des croisements contrôlés entre des plantes parmi la descendance desquelles ils ont sélectionnés les individus qui ont gardé les critères d'intérêts ; progressivement des variétés homogènes et stables ont été créés et diffusées. Dans le cas des arbres fruitiers les descendants de ces croisements contrôlés ont souvent été multipliés par propagation végétative ; c'est ainsi que des variétés modernes homogènes sont cultivées dans des vergers commerciaux et remplacent les variétés locales des vergers traditionnels. Cette sélection homogénéisante réduit drastiquement la diversité génétique. L'abandon de la culture d'un cultivar par les agriculteurs ou même par toute la société est la face drastique de la sélection. Cet abandon peut être la cause de l'extinction du cultivar lorsqu'aucun plan d'action de gestion des ressources phytogénétiques n'est mis en place.

II.1.2 Migration/diffusion

“In considering the distribution of organic beings over the face of the globe, the first great fact which strikes us is that neither the similarity nor the dissimilarity of the inhabitants of various regions can be accounted for by their climate and other physical conditions... The dissimilarity of inhabitants of different regions may be attributed to modification through natural selection... The degree of dissimilarity will depend on the migration of the more dominant forms of life from one region into another... Thus, the high importance of barriers comes into play by checking migration; as does time for the slow process of modification through natural selection”. Darwin (1859), p.181

La migration est l'échange d'individus ou de gamètes entre des populations d'une même espèce, elle permet les flux de gènes. En l'absence d'autres forces évolutives la migration homogénéise les fréquences alléliques entre les populations. La migration influence donc la distribution spatiale du polymorphisme intra-spécifique.

Chez les espèces cultivées, la structuration et la dynamique de la diversité génétique sont aussi influencées par les migrations humaines. La migration dépend d'une part des caractéristiques biologiques de l'espèce concernée et d'autre part des modalités humaines de sa reproduction et de sa dissémination ainsi que des interactions entre ces forces. A la suite de la domestication, l'Homme a fait circuler les espèces cultivées ce qui a brisé les barrières géographiques. Les espèces cultivées ont accompagné les grands voyageurs depuis des siècles (Egea and Garcia, 1975 ; Albala, 2009). Les anciennes routes de commerce ont été des passages pour la diffusion de différentes espèces cultivées d'une part dans l'ancien monde (Diamond and Bellwood, 2003) et puis entre l'ancien et le nouveau monde, à partir du 15^{ème} siècle (Ferrão and Castro, 2015). Ces vagues de diffusion successives ont permis l'introduction d'espèces cultivées dans de nouvelles régions; dans certains cas l'introduction a été réalisée dans des régions où des espèces sauvages apparentées étaient présentes. En conséquence, pour des espèces fruitières telles que le pommier, l'olivier ou l'amandier, des hybridations ont résulté de flux de gènes entre le taxon cultivé et les taxons spontanés devenus sympatriques ; ils ont ainsi contribué à l'évolution et à la diversification du taxon

cultivé (Cornille et al. 2012, 2013 ; Breton, Tersac and Bervillé, 2006 ; Haouane, 2012 ; Certal et al., 2002 ; Delplancke et al., 2012).

La biologie de reproduction et le mode de propagation adopté par les agriculteurs pour diffuser une espèce ou une variété limitent ou favorisent les flux de gènes. Le système de reproduction autogame et la propagation végétative limitent les flux de gènes, ce qui favorise une structuration génétique. Par contre, le système de reproduction allogame favorise les flux de gènes par pollinisation. Plus les flux de pollen sont importants entre des individus, plus l'homogénéisation génétique sera importante ; ceci conduit à une réduction de la différenciation génétique entre populations et dans une situation extrême les populations constituent un seul pool de gènes. La différenciation peut cependant être maintenue par les pratiques de sélection des agriculteurs. De plus, une espèce cultivée propagée par semis maintient plus de diversité intra-variétale qu'une espèce reconduite par clonage, sous la forme de boutures ou de greffes.

II.1.3 Dérive

La dérive génétique est un changement aléatoire des fréquences alléliques dû à un échantillonnage aléatoire des allèles à chaque génération dans une population de taille finie (Wright, 1931). La dérive mène à l'élimination ou à la fixation d'un allèle pour tous les individus d'une population. La dérive se répète à chaque génération, elle a tendance à réduire la diversité ; elle peut être contrebalancée par des mutations.

La domestication a été à l'origine d'importants goulots d'étranglement qui ont pu conduire à des dérives génétiques. Selon Hodgkin et al. (2007), les pratiques de sélection des agriculteurs peuvent causer une dérive génétique, sur les caractères non sélectionnés, dans les populations des espèces cultivées. L'effet de la dérive est négativement corrélé à la surface cultivée et à la quantité de graines ou de plants utilisée pour constituer chaque génération ou chaque nouveau

verger. Dans les agrosystèmes traditionnels la migration peut permettre de compenser les effets de la dérive et participer au maintien de la diversité génétique chez les espèces cultivées.

II.3 Facteurs culturels et diversité des plantes cultivées

La coexistence de l'Homme avec son environnement lui a apporté des connaissances qui lui ont permis d'utiliser les plantes présentant pour lui un intérêt. Les plantes cultivées sont aussi bien des objets biologiques que sociaux car elles sont héritées de génération en génération, intégrés dans les systèmes locaux de classification et de nomenclature, associés à des usages ou des rites, ou encore échangées entre voisins, amis, parents et alliés (Brush and Perales, 2007 ; Leclerc and Coppens, 2012; Hmimsa, Aumeeruddy-Thomas, and Ater, 2012 ; Labeyrie et al., 2016). Des études dans les agrosystèmes traditionnels ont montré l'intérêt des agriculteurs pour garder une diversité variétale; les agriculteurs définissent un cultivar par des critères identifiables et distinctifs qu'ils cherchent à exploiter et à maintenir (Elias et al., 2001; Barnaud et al., 2007). Chaque cultivar est ainsi nommé pour permettre la communication à son sujet et faciliter sa transmission (Boster, 1986).

Pour comprendre l'influence de l'Homme sur la structuration et la dynamique de la diversité des plantes cultivées plusieurs études ont cherché à identifier l'influence des aspects culturels sur les forces évolutives. Des études menées sur le maïs au Mexique ont montré que la sélection des semences est une pratique qui dépend de la perception et la représentation de la diversité de l'espèce cultivée (Pressoir and Berthaud, 2004 ; Perales, Benz, and Brush, 2005). Les travaux sur les perceptions et les représentations, les usages et les pratiques autour de différentes espèces cultivées se sont appuyés sur des nomenclatures et des classifications locales (Mekbib, 2007 ; Hmimsa, Aumeeruddy-Thomas and Ater, 2012). L'étude menée par Hmimsa, Aumeeruddy-Thomas and Ater (2012) a permis de mettre en évidence une diversité

variétale importante en étudiant la nomenclature du figuier (*Ficus carica*) dans le Rif du Maroc. Cette étude a démontré aussi que les agriculteurs gardent les variétés transmises des parents même si elles sont moins rentables. Le maintien de la diversité intra-spécifique du figuier dans le Rif est expliqué par des facteurs socio-économiques, culturels et religieux ; qu'elle soit fraîche ou séchée, la figue occupe une place centrale dans l'échange de cadeau entre amis et proches, surtout pendant les fêtes religieuses (Hmimsa, Aumeeruddy-Thomas and Ater, 2012).

III. Problématique de la thèse

III.1 Modèle d'étude: cadre bio géographique de la thèse

III.1.1 L'amandier

Biologie de la reproduction

L'amandier (*Prunus dulcis* (Mill.) D.A. Webb. syn. *Prunus amygdalus* Batsch., *Amygdalus communis* L., *Amygdalus dulcis* Mill.) est une espèce diploïde ($2n=16$). La pollinisation de l'amandier est entomophile, effectuée principalement par les abeilles mellifères (*Apis mellifera*). C'est une espèce à reproduction sexuée régulée par une auto-incompatibilité gamétophytique (GSI) (Socias i Company et al., 1976) ; ce mécanisme est régi par un seul locus S, polymorphe, comprenant au moins deux gènes liés, l'un spécifiquement exprimé dans le pistil et l'autre dans le pollen (Kao and Tsukamoto, 2004). Le système d'auto-incompatibilité est basé sur la reconnaissance de l'allèle lors de la croissance du tube pollinique ; la croissance du tube est arrêtée dans le style chaque fois que l'allèle S simple exprimé dans le pollen haploïde correspond à l'un des deux haplotypes S exprimés dans le tissu diploïde du pistil. Fernandez i Marti et al. (2009) ont démontré la présence de génotypes auto-compatibles qui ne possèdent pas l'allèle d'auto-compatibilité (Sf), ce qui les a conduit à suggérer la présence de gènes modificateurs en dehors de la région codante. La présence de

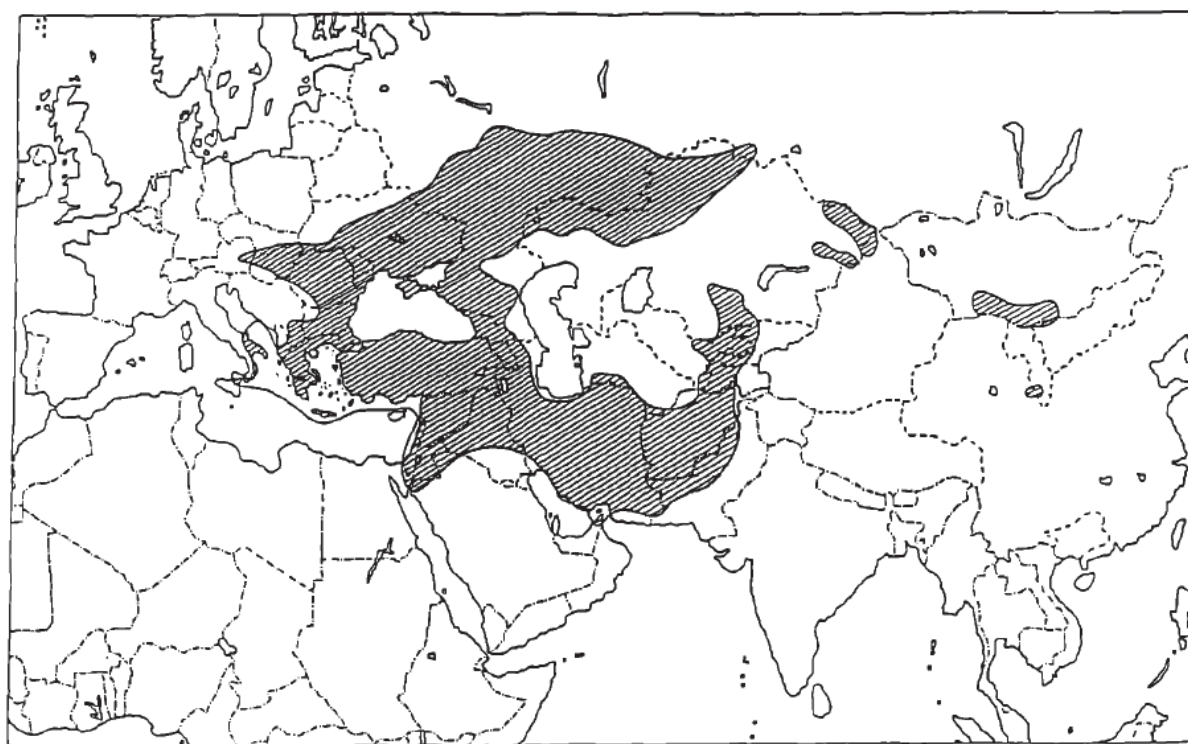


Figure 1. Distribution naturelle du genre *Amygdalus* L. selon Browicz and Zohary 1996.

gènes modificateurs confirme la proposition de Socias i Company (1990) selon laquelle l'amandier est une espèce auto-incompatible avec un arrière-plan génétique de pseudo-compatibilité. Sur ce fond, un seul allèle Sf pourrait casser le système d'auto-incompatibilité, mais probablement en interagissant avec ce contexte de pseudo-auto-compatibilité. Cette interaction a été démontrée par l'effet de deux QTL liés à l'expression de l'auto-compatibilité chez des géotypes ne montrant pas la présence de l'allèle Sf, donc théoriquement auto-incompatible (Fernández Martí et al., 2011).

La plupart des cultivars d'amandier sont auto-incompatibles (Socia i Company, 1990). Pour s'assurer que les fleurs soient efficacement pollinisées pour atteindre une production économiquement acceptable (Kester and Griggs, 1959), une pollinisation croisée de cultivars inter-compatibles et fleurissants de manière synchrone est nécessaire.

Histoire évolutive

La famille des Rosaceae contient de nombreuses espèces fruitières majeures, notamment dans le genre *Prunus*, dont les 430 espèces incluent les prunes, cerises, abricots, pêches, nectarines et amandes. L'amandier (*P. dulcis*) et le pêcher (*P. persica*) appartiennent au sous-genre *Amygdalus*, respectivement à ses sections *Amygdalus* et *Persica*. Ces deux sections auraient divergé il y a cinq millions ans (Delplancke et al., 2016), en relation avec la formation de chaînes de montagnes en Asie centrale (Sun et al., 1998). La section *Amygdalus* s'est différenciée à l'ouest de ce massif, dans les montagnes irano-touraniennes, sous des conditions arides (Browicz, 1989). Une trentaine d'espèce s'est formée par adaptations locales allopatriques; ces espèces sont réparties dans l'hémisphère nord sur une vaste zone géographique allant de l'Asie occidentale (ouest de la Chine) jusqu'au centre du bassin méditerranéen (Italie) (Figure 1). Quelques populations isolées ont été identifiées en Mongolie, en Chine centrale et au centre de l'Espagne. Bien avant la domestication, ces

espèces d'amandiers sauvages avaient déjà une importance économique et culturelle. En effet, des études archéologiques indiquent que les amandes de certaines de ces espèces sauvages ont été consommées et cultivées par les communautés méditerranéennes préhistoriques (Martinoli, and Jacomet. 2004 ; Willcox, Fornite, and Herveux 2007).

Si la plupart des auteurs conviennent de fixer la domestication de l'amandier au troisième millénaire avant notre ère, le processus en est controversé. Certaines études, soutiennent un scénario d'hybridation préalable entre différentes espèces de la section *Amygdalus* présentes dans les montagnes irano-touraniennes, et réfutent l'existence de populations sauvages de *P. dulcis* (Grasselly and Crossa-Raynaud, 1980). D'autres études soutiennent un scénario de domestication à partir de populations sauvages de *P. dulcis* au Levant, en réponse à la sélection pour des amandes douces (Browicz and Zohary, 1996). Les différentes recherches s'accordent sur un processus diffus dans l'espace et dans le temps (Weiss, Kislev and Hartmann, 2006 ; , Willcox, Buxo and Herveux, 2009), avec une culture prolongée de formes non domestiquées et une hybridation répétée avec des espèces sauvages locales (Delplancke et al., 2012 ; 2016).

En l'absence de reproduction végétative, une difficulté majeure de la domestication des arbres tient à leur allogamie qui limite fortement la transmission des caractères sélectionnés aux descendance obtenues par semis (Zohary and Spiegel-Roy, 1975). Dans le cas particulier de l'amandier, le facteur crucial qui en a permis la domestication a été la dominance de l'allèle de l'amandon doux sur l'allèle récessif de l'amandon amer, présent dans les formes sauvages. Cette dominance simple a permis la transmission par semis de ce caractère essentiel à une part assez importante de la descendance des individus sélectionnés. Les découvertes archéobotaniques puis des registres historiques ont montré que l'amandier domestiqué a été diffusé par graines transportées le long des anciennes routes de commerce (e.g. la Route de la

soie ; Albala, 2009). La région de départ de la diffusion, le massif irano-touranien ou le Levant, doit être envisagé selon le scénario de domestication défendu. A partir du Levant, l'amandier cultivé aurait été introduit dans l'ensemble du Bassin Méditerranéen, successivement par les Phéniciens, les Grecs, les Romains et les Arabes (Egea and García, 1975; Jaouani, 1976 ; Laghezali, 1985). Dans ce cas, la majorité de l'amandier cultivé dans le Bassin Méditerranéen provient d'un seul pool de gènes, comme cela a été démontré par Delplancke et al. (2013) avec des marqueurs SSR chloroplastiques. Suite à la diffusion par l'homme, la présence en sympatrie du taxon cultivé et de taxons apparentés spontanés a conduit à des hybridations entre les différents compartiments. Ces flux de gènes contribuent à l'évolution au fil de la domestication de l'amandier cultivé. Ils ont notamment permis l'introggression du gène d'auto-compatibilité, depuis des populations italiennes de *P. webbii* (Grasselly, 1976 ; Certal et al., 2002). De même, des flux de gènes et des hybrides entre *P. dulcis* et *P. orientalis* ont été identifiés au Proche-Orient (Delplancke et al., 2012). Enfin, l'amandier a été introduit dans le Nouveau Monde par les colons européens, à partir d'un nombre limité de cultivars européens (Wood, 1925).

En l'espace de deux mille ans, l'amandier, toujours propagé par semis, a fait l'objet d'une sélection naturelle dans divers milieux, ce qui a permis à cette espèce d'évoluer comme une espèce forestière (Grasselly and Crossa-Raynaud, 1980). Sélections artificielle et naturelle ont abouti à la différenciation d'écotypes bien définis devenant des cultivars locaux ou traditionnels (Grasselly, 1976 ; Grasselly and Crossa-Raynaud, 1980). Ainsi, l'amandier cultivé dans le Bassin Méditerranéen s'est différencié en plusieurs pools de gènes distribués géographiquement, identifiés par Delplancke et al. (2013) en utilisant des marqueurs SSR nucléaires. A partir du 19^{ème} siècle, les différentes sélections locales d'amandier ont été propagées par greffage pour augmenter et uniformiser la production dans différents pays du Bassin Méditerranéen (France, Espagne, Tunisie et Italie) (Grasselly and Crossa-Raynaud,

1980). Par ailleurs, les amandiers introduits en Amérique ont aussi été sélectionnés et propagés par greffage pour répondre aux besoins croissants d'une productivité commerciale élevée. Ces pratiques de sélection ont fait émerger un pool de gènes américain différencié du pool génétique européen, identifié par Fernández i Martí et al. (2015) en utilisant des marqueurs SSR nucléaires. Les programmes de sélection moderne en Méditerranée et aux Etats-Unis ont ainsi produit des variétés commerciales propagées par greffage ce qui a conduit à une uniformité de la production. Certaines de ces variétés modernes ont été introduites dans différents pays du Bassin Méditerranéen et ont remplacé des cultivars locaux dans les vergers commerciaux ; cette pratique porte atteinte à la dynamique évolutive de l'amandier cultivé. L'effet de l'introduction d'une variété propagée par greffage sur la structuration et la diversité génétique sera discuté plus en détail dans le troisième chapitre.

III.1.2 Le Liban

«Accueillant comme ses rades, aventureux comme sa Méditerranée, sonore comme ses torrents, s'offrant comme ses sources, miroitant comme ses étoiles; robuste, aimable comme ses collines, âpre et démunie comme ses terres escarpées, le Liban « fille de la géographie » s'est lentement, habilement façonné au gré de l'histoire.» André Chédid (1969)

Entre 1500 et 300 av. J.-C., les Phéniciens vivaient dans le Liban actuel et la Syrie côtière, ils ont dirigé certains des premiers grands réseaux commerciaux de la Méditerranée et dominaient la mer avec les Grecs (Pottier, 1920). Certains naviguèrent jusqu'aux îles britanniques et beaucoup d'entre eux établirent des colonies en Afrique du Nord, en Espagne, en Sicile et en Sardaigne. Ce fut l'un des premiers liens culturels et aussi de diffusion de l'amandier entre le Moyen-Orient et les autres rives du Bassin Méditerranéen. Les travaux réalisés en génétique par Delplancke et al. (2013) ont souligné i) la présence de différents pools de gènes d'amandier cultivé dans le Bassin Méditerranéen; ii) l'importance du Levant comme centre de diversification de *P. dulcis*, considéré comme un réservoir de diversité génétique de l'amandier cultivé. Au Liban, on trouve encore des populations spontanées ou

férales de *P. dulcis*, de *P. orientalis* et de *P. spartioides* (Post, 1932 ; Mouterde, 1966 ; Talhouk et al., 2000 ; Chalak and Hamadeh, 2014) dont certaines ont contribué au pool génétique de l'amandier cultivé ; par exemple, Delplancke et al. (2012) ont montré l'existence de flux de gènes entre *P. orientalis* et *P. dulcis*. Chalak et al. (2006) ont réalisé une première étude de la diversité génétique de l'amandier cultivé au Liban à partir d'un échantillon de 60 individus collectés dans les vergers familiaux ; leurs résultats montrent l'existence d'une grande diversité dans ces vergers.

Des programmes de conservation ont été développés, des efforts de reforestation à partir de *P. dulcis* et *P. orientalis* ont été réalisés entre 1960 et 1975 dans le cadre du programme 'Plan vert' du Ministère de l'Agriculture en installant des parcelles réservoirs en particulier dans la Bekaa (Talhouk et al., 2000). Toutefois, le Liban n'a pas élaboré de programme d'amélioration ; les variétés améliorées sont toujours introduites à partir de différents pays et elles sont propagées par greffage dans des vergers commerciaux. Ceux-ci sont constitués d'un nombre limité de variétés améliorées. Ces cultures commerciales sont connues pour la réduction de leur diversité génétique et leur faible capacité adaptative face aux changements climatiques et aux ravageurs. Après avoir longtemps constitué la base du verger libanais, le matériel local est donc progressivement remplacé par de nouvelles variétés provenant d'autres pays. Le Liban héberge une diversité intra-spécifique importante, mais l'évaluation des conséquences à long terme de ces nouvelles pratiques de reconduction des vergers constitue un enjeu d'importance pour comprendre la dynamique actuelle de la diversité variétale de l'amandier cultivé.

En plus de son intérêt fondamental, la compréhension de la dynamique de la diversité inter et intra-variétale constitue un enjeu considérable pour la gestion et la préservation des espèces

cultivées dans un contexte de changements démographiques, climatiques, agricoles, et politiques.

III.2 Questions de recherche et démarche

Dans ce travail je me suis intéressée aux questions suivantes :

- Quel est le degré de parenté entre les cultivars libanais et ceux d'autres pays méditerranéens? Quel est l'effet des facteurs historiques et récents sur la dynamique de diffusion de ces cultivars d'amandier? (Chapitre 1)
- Quelle est l'importance des facteurs culturels dans cette structuration ? (Chapitre 2)
- Quel est l'effet du mode de propagation et de la dynamique de diffusion sur la diversité variétale? (Chapitre 3)

La démarche suivie s'inscrit dans le domaine des interactions bio-culturelles, faisant appel à la biologie des populations et à l'anthropologie. L'analyse de la diversité génétique des amandiers du Liban permettra de comprendre la dynamique de structuration du matériel cultivé ainsi que les mécanismes d'évolution de cette espèce dans les agroécosystèmes traditionnels. L'approche comparative des variétés d'amandier du Liban avec les variétés provenant d'autres pays méditerranéens permettra d'identifier les relations génétiques entre ces variétés et de mettre en évidence d'éventuelles particularités du matériel libanais. Cette approche comparative permettra aussi d'identifier les flux de gènes récents et de mieux comprendre la dynamique évolutive de cette espèce cultivée. L'approche anthropologique est basée sur un suivi d'enquêtes et d'observations dans les vergers, des listes libres et une approche sémantique.

III.3 Plan de la thèse

Après une introduction générale cette thèse est organisée en trois chapitres qui sont rédigés sous forme de manuscrits d'articles:

Chapitre 1, Impact of ancient and modern human-mediated gene flow on the genetic structure of cultivated almonds in the Mediterranean Basin. Nous avons évalué la diversification génétique de l'amandier cultivé dans le Bassin Méditerranéen et tenté de différencier l'impact des processus humains anciens des processus plus récents sur la structure de la diversité génétique. Cette étude nous a permis de relier une série d'événements historiques qui ont conduit à la diffusion continue de l'amandier cultivée depuis sa domestication avec la structure de sa diversité génétique.

Chapitre 2, Taxinomie locale et caractérisation de la diversité de l'amandier cultivée au Liban. Nous avons exploré la variabilité des connaissances locales autour de l'amandier cultivé au Liban : la diversité nommée, la nomenclature, les critères de classification à différentes échelles. Cette étude nous a permis de proposer une taxinomie locale pour l'amandier au Liban.

Chapitre 3, Evolution of almond genetic diversity and farmer practices in Lebanon: impacts of the diffusion of a graft-propagated cultivar in a traditional system based on seed-propagation. Nous avons évalué l'impact des deux modes de propagation, sexuée et clonale, sur la structure et la dynamique de la diversité génétique des deux cultivars phares identifiés dans le chapitre 2. Cette étude a permis d'évaluer l'impact de différentes pratiques de propagation sur la diversité génotypique et génétique de chaque cultivar.

Enfin, la dernière partie propose une discussion générale et des perspectives ; les principaux résultats sont discutés et mis en relation par un regard croisé génétique/ anthropologie pour mettre en évidence les interactions entre facteurs biologiques et facteurs anthropiques qui

influencent les forces évolutives et le niveau de diversité génétique inter et intra-variétale. Quelques implications saillantes de nos résultats sont présentées pour la conservation et la gestion durable de l'amandier au Liban.

Bibliographie

Abbo, S. *et al.* (2013) 'Plant domestication versus crop evolution: a conceptual framework for cereals and grain legumes', *Trends in Plant Science*, 19(6), pp. 1360–1385. doi: 10.1016/j.tplants.2013.12.002.

Albala, K. (2009) 'Adaptation of Ideas from West to East', *Petits Propos Culinaire*, 88, pp. 19–34.

Barnaud, A. *et al.* (2007) 'Local genetic diversity of sorghum in a village in northern Cameroon: Structure and dynamics of landraces', *Theoretical and Applied Genetics*, 114(2), pp. 237–248. doi: 10.1007/s00122-006-0426-8.

Boster, J. S. (1986) 'Exchange of Varieties and Information between Aguaruna Manioc Cultivators', *American Anthropologist*, 88(2), pp. 428–436. doi: 10.1525/aa.1986.88.2.02a00100.

Breton, C., Tersac, M. and Bervillé, A. (2006) 'Genetic diversity and gene flow between the wild olive (oleaster, *Olea europaea* L.) and the olive: several Plio-Pleistocene refuge zones in the Mediterranean basin suggested by simple sequence repeats analysis', *Journal of Biogeography*. Wiley/Blackwell (10.1111), 33(11), pp. 1916–1928. doi: 10.1111/j.1365-2699.2006.01544.x.

Browicz, K. and Zohary, D. (1996) 'The genus *Amygdalus* L. (Rosaceae): Species relationships, distribution and evolution under domestication', *Genetic Resources and Crop Evolution*, 43(3), pp. 229–247. doi: 10.1007/BF00123275.

Brush, S. B. and Perales, H. R. (2007) 'A maize landscape: Ethnicity and agro-biodiversity in Chiapas Mexico', *Agriculture, Ecosystems and Environment*, 121(3), pp. 211–221. doi: 10.1016/j.agee.2006.12.018.

Certal, A. C. *et al.* (2002) 'Structural and molecular analysis of self-incompatibility in almond (*Prunus dulcis*)', *Sexual Plant Reproduction*, 15(1), pp. 13–20. doi: 10.1007/s00497-002-0138-4.

Chalak, L. *et al.* (2006) 'Preliminary Characterization of Cultivated Almonds (*Prunus dulcis* L.) in Lebanon by Morphological Traits and Microsatellite Markers', *Biologia Tunisie*, 4, p. 0.

Chalak, L. and Hamadeh, B. (2014) 'Almond wild relatives in Lebanon: distribution, uses and main threats', *II International Symposium on Wild Relatives*. Available at: https://www.actahort.org/books/1074/1074_5.htm

Charles Grasselly, E. D. (2016) '1978 / Lauranne : la première variété autofertile d'amandier commercialisée', *INRA Science & Impact*. Available at: <http://www.bordeaux->

aquitaine.inra.fr/Toutes-les-actualites/70-ans-Lauranne-premiere-variete-amandier-autofertile-commercialisee (Accessed: 11 July 2018).

Choueiri, E., Verdin, E. Danet, J. L., Jreijiri, F., El Zammar, S., Salard, P., Bové, J.M., Garnier, M. (2003) *A phytoplasma disease of almond in Lebanon [2003]*, *AGRIS: International Information System for the Agricultural Science and Technology*. Available at: <http://agris.fao.org/agris-search/search.do?recordID=QC2003001784> (Accessed: 11 July 2018).

Cornille, A. *et al.* (2012) ‘New Insight into the History of Domesticated Apple: Secondary Contribution of the European Wild Apple to the Genome of Cultivated Varieties’, *PLoS Genetics*. Edited by R. Mauricio. Public Library of Science, 8(5), p. e1002703. doi: 10.1371/journal.pgen.1002703.

Cornille, A., Gladieux, P. and Giraud, T. (2013) ‘Crop-to-wild gene flow and spatial genetic structure in the closest wild relatives of the cultivated apple’, *Evolutionary Applications*, 6, pp. 737–748. doi: 10.1111/eva.12059.

Darwin, C. (1859) ‘The Origin of The Species’. pp502 Available at: http://www.vliz.be/docs/Zeeijfers/Origin_of_Species.pdf.

Darwin, C. (1868). ‘The Variation of Animals and Plants Under Domestication’. Volume 1. pp 411;

Delplancke, M. *et al.* (2012) ‘Gene flow among wild and domesticated almond species: Insights from chloroplast and nuclear markers’, *Evolutionary Applications*, 5(4), pp. 317–329. doi: 10.1111/j.1752-4571.2011.00223.x.

Delplancke, M. *et al.* (2013) ‘Evolutionary history of almond tree domestication in the Mediterranean basin’, *Molecular Ecology*. Wiley/Blackwell (10.1111), 22(4), pp. 1092–1104. doi: 10.1111/mec.12129.

Delplancke, M. *et al.* (2013) ‘Evolutionary history of almond tree domestication in the Mediterranean basin’, *Molecular ecology*. Wiley Online Library, 22(4), pp. 1092–1104.

Delplancke, M. *et al.* (2016) ‘Combining conservative and variable markers to infer the evolutionary history of *Prunus* subgen. *Amygdalus* s.l. under domestication’, *Genetic Resources and Crop Evolution*, 63(2), pp. 221–234. doi: 10.1007/s10722-015-0242-6.

Diamond, J. (2002) ‘Evolution, consequences and future of plant and animal domestication’, *Nature*, 418, pp. 700–707. doi: 10.1038/nature01019.

Diamond, J. and Bellwood, P. (2003) ‘Farmers and Their Languages: The First Expansions’, *Science*, 300(597), p. 597. Available at: www.sciencemag.org (Accessed: 10 July 2018).

Elias, M. *et al.* (2001) ‘Traditional management of cassava morphological and genetic diversity by the makushi amerindians (Guyana, South America): Perspectives for on-farm conservation of crop genetic resources’, *Euphytica*, 120(1), pp. 143–157. doi: 10.1023/A:1017501017031.

FAO 1996. *LEIPZIG DECLARATION ON CONSERVATION AND SUSTAINABLE UTILIZATION OF PLANT GENETIC RESOURCES FOR FOOD AND AGRICULTURE*. Available at: <http://www.fao.org/FOCUS/E/96/06/more/declar-e.htm>.

- Fernandez i Marti, A. *et al.* (2009) 'Genetic Diversity in Spanish and Foreign Almond Germplasm Assessed by Molecular Characterization with Simple Sequence Repeats', *Journal of American Society Horticulture Science*, 134(5), pp. 535–542. Available at: <http://journal.ashspublications.org/content/134/5/535.short>.
- Fernández i Martí, A. *et al.* (2015) 'Molecular analyses of evolution and population structure in a worldwide almond [*Prunus dulcis* (Mill.) D.A. Webb syn. *P. amygdalus* Batsch] pool assessed by microsatellite markers', *Genetic Resources and Crop Evolution*, 62(2), pp. 205–219. doi: 10.1007/s10722-014-0146-x.
- Fernández Martí, À. *et al.* (2011) 'Identification of quantitative trait loci associated with self-compatibility in a *Prunus* species', *Tree Genetics & Genomes*, 7, pp. 629–639. doi: 10.1007/s11295-010-0362-2.
- Ferrão, J. E. M. and Castro, X. de (2015) *Le voyage des plantes & les Grandes Découvertes (XVe-XVIIe siècles. Chandeigne. Available at: <https://editionschandeigne.fr/livre/le-voyage-des-plantes-les-grandes-decouvertes/>*
- Fuller, D. Q. (2007) 'Contrasting Patterns in Crop Domestication and Domestication Rates: Recent Archaeobotanical Insights from the Old World', *Annals of Botany*. Oxford University Press, 100(5), pp. 903–924. doi: 10.1093/aob/mcm048.
- Glémin, S. and Bataillon, T. (2009) 'A comparative view of the evolution of grasses under domestication: Tansley review', *New Phytologist*, 183(2), pp. 273–290. doi: 10.1111/j.1469-8137.2009.02884.x.
- Gouraud, S. du général (1921) 'La culture des arbres fruitiers en Syrie et Cilicie', *Revue de botanique appliquée et d'agriculture coloniale*, 1, pp. 129–148. doi: <https://doi.org/10.3406/jatba.1921.4001>.
- Grasselly, C. (1976) 'Les espèces sauvages d'amandier', *Options méditerranéennes I*.
- Grasselly, C. and Crossa-Raynaud, P. (1980) *L'amandier*. Maisonneuv.
- Haouane, H. (2012) *Origines , domestication et diversification variétale chez l ' olivier (Olea europaea L .) à l ' ouest de la Méditerranée*.
- Harlan, J. R. (1992) 'Crops and man.', *Crops and man*. American Society of Agronomy, (Ed. 2). Available at: <https://www.cabdirect.org/cabdirect/abstract/19931634635>.
- Hmimsa, Y., Aumeeruddy-Thomas, Y. and Ater, M. (2012) 'Vernacular taxonomy, classification and varietal diversity of fig (*Ficus carica* L.) among Jbala cultivators in Northern Morocco', *Human Ecology*, 40(2), pp. 301–313. doi: 10.1007/s10745-012-9471-x.
- Hodgkin, T. *et al.* (2007) 'Seed systems and crop genetic diversity in agroecosystems', in *Managing biodiversity in agricultural ecosystems*, pp. 77–116. doi: 10.7312/jarv13648-006.
- Jaouani, A. (1976) 'La culture de l'amandier en Tunisie', *Cahiers Options Méditerranéennes*.
- Kao, T. and Tsukamoto, T. (2004) 'The molecular and genetic bases of S-RNase-based self-incompatibility.', *The Plant cell*. American Society of Plant Biologists, 16 Suppl(Suppl), pp. S72-83. doi: 10.1105/tpc.016154.
- Labeyrie, V. *et al.* (2016) 'Seed exchange networks, ethnicity, and sorghum diversity',

Proceedings of the National Academy of Sciences, 113(1), pp. 98–103. doi: 10.1073/pnas.1513238112.

Laghezali, M. (1985) 'L'amandier au Maroc', *Options Méditerranéennes*.

Leclerc, C. and Coppens, G. (2012) 'Social Organization of Crop Genetic Diversity. The $G \times E \times S$ Interaction Model', *Diversity*, 4, pp. 1–32. doi: 10.3390/d4010001.

Mariette, S. *et al.* (2010) 'Population structure and genetic bottleneck in sweet cherry estimated with SSRs and the gametophytic self-incompatibility locus', *BMC Genetics*, 11, p. 77. doi: 10.1186/1471-2156-11-77.

McKey, D. *et al.* (2012) 'Revue d'ethnoécologie', 1(2012), pp. 1–25.

Mekbib, F. (2007) 'Infra-specific folk taxonomy in sorghum (*Sorghum bicolor* (L.) Moench) in Ethiopia: folk nomenclature, classification, and criteria', 18, pp. 1–18. doi: 10.1186/1746-4269-3-38.

Meyer, R. S., DuVal, A. E. and Jensen, H. R. (2012) 'Patterns and processes in crop domestication: an historical review and quantitative analysis of 203 global food crops.', *The New Phytologist*, 196(1), pp. 29–48. doi: 10.1111/j.1469-8137.2012.04253.x.

Meyer, R. S. and Purugganan, M. D. (2013) 'Evolution of crop species: Genetics of domestication and diversification', *Nature Reviews Genetics*, 14(12), pp. 840–852. doi: 10.1038/nrg3605.

Miller, A. J. and Gross, B. L. (2011) 'From forest to field: Perennial fruit crop domestication', *American Journal of Botany*, 98(9), pp. 1389–1414. doi: 10.3732/ajb.1000522.

Mouterde, P. (1966) *Nouvelle flore du Liban et de la Syrie. Tome I*. Imprimerie. Beirut, Lebanon.

Perales, H. R., Benz, B. F. and Brush, S. B. (2005) 'Maize diversity and ethnolinguistic diversity in Chiapas, Mexico', *Proceedings of the National Academy of Sciences*, 102(3), pp. 949–954. doi: 10.1073/pnas.0408701102.

Post, G. E. (1932) *Flora of Syria, Palestine, and Sinai*.

Pottier, E. (1920) 'C.Autran — Phéniciens. Essai de contribution à l'histoire antique de la Méditerranée.', *Syria. Archéologie, Art et histoire*. Persée - Portail des revues scientifiques en SHS, 1(1), pp. 329–332. Available at: https://www.persee.fr/doc/syria_0039-7946_1920_num_1_1_2851

Pressoir, G. and Berthaud, J. (2004) 'Patterns of population structure in maize landraces from the Central Valleys of Oaxaca in Mexico', *Heredity*, 92(2), pp. 88–94. doi: 10.1038/sj.hdy.6800387.

Purugganan, M. D. and Fuller, D. Q. (2009) 'The nature of selection during plant domestication', *Nature*, 457(7231), pp. 843–848. doi: 10.1038/nature07895.

Stockwell, C. A., Hendry, A. P. and Kinnison, M. T. (2003) 'Contemporary evolution meets conservation biology', *Trends in Ecology and Evolution*, 18(2), pp. 94–101. Available at: <http://tree.trends.com>

- Sun Donghuai *et al.* (1998) 'Magnetostatigraphy and palaeoclimatic significance of late Tertiary aeolian sequences in the Chinese Loess Plateau', *Geophysical Journal International*, 134(1), pp. 207–212. doi: 10.1046/j.1365-246X.1998.00553.x.
- Talhouk, S. N. *et al.* (2000) 'Phenotypic diversity and morphological characterization of *Amygdalus L.* species in Lebanon', *Genetic Resources and Crop Evolution*. Springer, 47(1), pp. 93–104.
- Tanno, K.-I. and Willcox, G. (2012) 'How Fast Was Wild Wheat Domesticated?', *Science*, 311(5769), p. 1886. doi: 10.1126/science.1124635.
- Weiss, E., Kislev, M. E. and Hartmann, A. (2006) 'Autonomous cultivation before domestication', *Science*, 312, pp. 1608–1610. doi: 10.1126/science.1127235.
- Willcox, G., Buxo, R. and Herveux, L. (2009) 'Late Pleistocene and early Holocene climate and the beginnings of cultivation in northern Syria', *The Holocene*. SAGE PublicationsSage UK: London, England, 19(1), pp. 151–158. doi: 10.1177/0959683608098961.
- Willcox, G., Fornite, S. and Herveux, L. (2007) 'Early Holocene cultivation before domestication in northern Syria', *Veghet Hist Archeobot*, 17, pp. 313–325. doi: 10.1007/s00334-007-0121-y.
- Wright, S. (1931) 'Evolution in Mendelian Populations.', *Genetics*, 16(2), pp. 97–159. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17246615>
- Zeder, M. A. *et al.* (2006) 'Documenting domestication: The intersection of genetics and archaeology', *Trends in Genetics*, 22(3), pp. 139–155. doi: 10.1016/j.tig.2006.01.007.
- Zohary, D. (2004) 'Unconscious selection and the evolution of domesticated plants', *Economic Botany*, 58(1), pp. 5–10. doi: 10.1663/0013-0001(2004)058[0005:USATEO]2.0.CO;2.
- Zohary, D. and Spiegel-Roy, P. (1975) 'Beginnings of Fruit Growing in the Old World', *Science*, 187(4174), pp. 319–327. doi: 10.1126/science.187.4174.319.

Chapitre 1

Impact of ancient and modern human-mediated gene flow on the genetic structure of cultivated almonds in the Mediterranean Basin

In preparation

Bariaa Hamadeh^{1,2*,3,4}, Geo Coppens d'Eeckenbrugge^{2,3}, Lamis Chalak⁵, Hélène I. Joly^{2,3}

1 Lebanese University, Doctoral School of Sciences and Technologies, Hadath, Lebanon.

2 CIRAD, UMR AGAP, F-34398 Montpellier, France.

3 AGAP, Univ. Montpellier, CIRAD, INRA, Montpellier SupAgro, Montpellier, France.

4 Université Montpellier, UMR CEFE, 34293 Montpellier cedex, France.

5 Lebanese University, Faculty of Agriculture, Dekwaneh, Lebanon.

* Present address

E-mail: bariaa.hamadeh@gmail.com; lamis.chalak@ul.edu.lb; geo.coppens@cirad.fr;
helene.joly@cirad.fr

Abstract

The organization of the genetic diversity of cultivated species is shaped by the effects of natural adaptive selection along with human selection and diffusion. Comprehending effects of different spatial and temporal human-drivers on the structure of the genetic diversity is therefore essential to assess the evolutionary dynamics of cultivated species and their continuous diversification. This study aims at deciphering the evolutionary processes related to human activities, involving exchanges and propagation practices, at different temporal and spatial scales.

Our results showed that recent human-driven effects increase the complexity of the evolutionary dynamics of cultivated almond. The traditional almond sampled from the Mediterranean Basin countries was distributed over five groups linked with different geographic regions. The spatial and temporal dynamic of human-mediated gene flow by seeds exchanged between geographically distinct groups, since domestication, has contributed to increase the genetic diversity of each group and reduced the genetic differentiation among them. The backward diffusion to the eastern part of the Mediterranean Basin was ensured by vegetative propagation, introgression was ensured later on by sexual propagation, evidenced by the presence of admixed individuals. This study allowed a better understanding of the historical dynamics of almond genetic diversity.

Keywords: cultivated almond, Mediterranean Basin, diversification, genetic diversity, gene flow, historical events

Background

The highly diverse genus *Prunus* (Rosaceae) includes about 430 stone-fruit species, including plums, cherries, apricots, peaches and almonds (Browicz and Zohary, 1996; Yazbek, 2010; Yazbek and Oh, 2013). The latter group belongs to the subgenus *Amygdalus* (L.) Focke section *Amygdalus*, whose type species is the cultivated almond (*Prunus dulcis* (Mill.) D.A. Webb. syn. *Prunus amygdalus* Batsch., *Amygdalus communis* L., *Amygdalys dulcis* Mill.), one of the oldest domesticated trees (Zohary, 2004). The current organization of its genetic diversity results from the whole succession of differentiation and hybridization events that have occurred from speciation within the genus until the relatively recent establishment of *P. dulcis* cultivars, through domestication.

Using SSR markers, Delplancke et al. (2016) estimated that section *Amygdalus* diverged from section *Persica* about five million years ago, which appears consistent with vicariance events resulting from the emergence of mountain ranges in Central Asia and subsequent climate changes (Sun et al., 1998). Almonds evolved on the western side of the Central Asian Massif in arid environments (Watkins, 1976), dispersing from western and central Asia to the Mediterranean Basin and differentiating into about 30 allopatric species (Vavilov, 1930; Zohary and Hopf, 1993; Gradziel, 2011; Yazbek and Oh, 2013). A few isolated populations were identified in Mongolia and Central China (Denisov, 1988) and in central Spain (Felipe and Socias I Company, 1977); their status as relics or late introductions is not clear.

Well before domestication, wild almond have played an important economic and cultural role. Archeological records indicate that wild almond nuts were consumed and cultivated by prehistoric Mediterranean communities (Martinoli and Jacomet, 2004; Willcox, Fornite, and Herveux, 2007). Usually, these almonds have bitter kernel, however individual trees producing sweet kernels exist in the wild (Vavilov, 1930; Denisov, 1988), providing a basis

for selection and domestication (Browicz and Zohary, 1996). While most authors agree in setting almond domestication in the third millennium BC, the process remains controversial, although the two proposed scenarios may not be mutually exclusive (Delplancke et al., 2016).

According to Browicz and Zohary (1996), and based on archaeological evidence, the pre-existing wild *P. dulcis* had extended, from Iran and eastern Turkey, into Syria, Lebanon and Jordan where it was subject to selection for sweet kernels. This so-called “Levant scenario”, giving a primary role to *P. dulcis*, is supported by the study of Delplancke et al. (2013) using chloroplastic and nuclear microsatellites. In contrast, according to Grasselly and Crossa-Raynaud (1980), several wild *Prunus* species from the Irano-Tauranian Mountains would have contributed to the domestication process, giving birth to *P. dulcis* as a cultigen. This “hybrid scenario” scenario is supported by the studies of Socias i Company and Felipe (1988), Kester, Gradziel and Grasselly (1991), and by recent works based on nuclear microsatellites and sequence data (Zeinalabedini et al., 2010; Delplancke et al., 2012; Delplancke et al., 2016). As to the center of origin, different studies support a geographically diffuse event for almond domestication (Weiss, Kislev and Hartmann, 2006; Willcox, Buxo and Herveux, 2009), involving the prolonged cultivation of exotic undomesticated or incipiently domesticated forms and their hybridization with local wild species (Zeder et al., 2006; Miller and Gross, 2011; Delplancke et al., 2016).

Either from the Irano-Tauranian Mountains through the Levant or directly from the Levant, domesticated almond seeds diffused along the ancient trade networks, including the ancient Silk Route. The presence of the cultivated almond in the eastern Mediterranean Basin is attested by archeological records from the second millennium BC (Zohary and Hopf, 1993), and supported by genetic analyses based on cpDNA SSR markers (Delplancke et al., 2013). This westward dissemination process started with Phoenicians and Greeks sea traders (Egea and Garcia, 1975). Almond was cultivated in Egypt before 1300 BC (Zohary and Hopf,

1993), and introduced by vessels into Greek islands around 350 BC (Albala, 2009). Almond was then introduced to all convenient areas of the Mediterranean Basin including, Italy, southern France, Spain, Portugal, and North Africa. More recently, Arab conquests in the 15th century introduced cultivated almond into Tunisia (Jaouani, 1976) and Morocco (Laghezali, 1985), then into Spain and Portugal (Egea and García, 1975).

For two thousand years, well-defined ecotypes and local cultivars have continued their evolution in the Mediterranean Basin through local selection and seed propagation (Grasselly, 1976; Grasselly and Crossa-Raynaud, 1980). Thus, the Mediterranean almond differentiated into gene pools (Delplancke et al., 2013). With the expansion of almond cultivation, new hybridizations have occurred between local cultivated almond and wild relatives. Thus, substantial introgressions were identified, involving *P. webbii* (Spach) Vierh. in Italy (Grasselly, 1976; Godini 1979, 2000; Certal et al., 2002; Socias i Company, 2004) and *P. orientalis* (Duhamel) in Lebanon and Syria (Delplancke et al., 2012). *P. webbii* is the likely source of a self-compatibility allele into sympatric cultivated almond (predominantly self-incompatible) (Grasselly, 1976 ; Godini, 2000; Certal et al., 2002) while *P. orientalis* contributed to the high genetic diversity of *P. dulcis* sympatric populations (Delplancke et al., 2012).

From the 19th century, almond grafting was practiced in the western Mediterranean Basin (France, Spain, Tunisia and Italy) to propagate local cultivars and improve almond productivity and homogeneity (Grasselly and Crossa-Raynaud, 1980). By the 20th century, improvement programs started in France selecting two cross-compatible cultivars ‘Ferragnès’ and ‘Ferraduel’ (Grasselly, 1976), which became the basis for new orchards in France and diffused to the rest of the Mediterranean regions. Later on, many countries including Tunisia, Italy and Spain developed their own selection programs (Grasselly, 1976). In 1974, these countries fostered cooperation among Mediterranean countries, establishing the G.R.E.M.P.A.

(Groupe de Recherches et d'Etudes Méditerranéen pour l'Amandier) (Grasselly and Crossa-Raynaud, 1980) and strengthening international exchanges in Europe, North Africa, Near East, America and Australia (Vargas, 2014). Modern vegetatively propagated cultivars were established in new commercial orchards, replacing many local cultivars in traditional orchards across the Mediterranean Basin. This has led to more uniformity, while affecting severely the dynamics of almond evolution. Nevertheless, almond culture systems based mainly on seed propagation persist in different Mediterranean countries, as in Mallorca, Greece, Turkey, Morocco, Syria, Iran (Gradziel, 2017), and Lebanon (Chalak et al., 2006). Such traditional orchards still represent the original local almond gene pools in the Mediterranean Basin.

In summary, hybridization events, involving both wild and cultivated almonds have modelled almond evolutionary history. Almond is mostly self-incompatible; it requires cross-pollination, which promotes genetic variability and adaptability to various environments (Socias i Company et al., in ed. Badenes and Byrne 2012). Human-driven migrations have contributed to the diffusion of cultivated almond (Gradziel, 2011). In the last decade, almond vegetative propagation has strongly reinforced human-driven effects on the genetic diversity of almond (Grassely and Crossa-Raynaud, 1980). As compared to other Mediterranean fruit trees (e.g. olive, fig and grape), this process is much more recent in the cultivated almond. Thus, this species provides a very interesting model for analyzing the effects of ancient and recent human-driven processes on the structure of the genetic diversity of cultivated fruit trees.

In the last decade, genetic studies have investigated the history of almond cultivation, and more particularly the geography of domestication events and the contribution of gene flow between wild and domesticated forms. They have highlighted the central role of the Middle East in the almond early dissemination, the importance of this region as a hotspot of almond genetic diversity, and the importance of hybridization between sympatric cultivated and wild

almonds. The genetic diversity of local cultivars, in Tunisia (Gouta et al., 2010), Morocco (Elhamzaoui et al., 2012), Italy (Distefano et al., 2013), and Spain (Fernandez i Marti et al., 2009), was characterized in comparison to accessions from collections, using nuclear microsatellites. Halász et al. (2016) observed the link between genetic diversity and mating system, a reduction of the genetic diversity in self-compatible Hungarian cultivars. However, to our knowledge, no study has investigated the effect of the human-drivers (e.g. different introductions) on the diversity and differentiation within cultivated almonds, except for our own study on the two most common cultivars in Lebanon (Hamadeh *et al.*, 2018).

In Lebanon, two sympatric wild relative species are still present next to seed propagated local cultivars and graft-propagated introduced cultivars. The first introduced cultivars received vernacular names such as ‘Halwani’, ‘Nahali’ from local farmers while others have conserved the name of their country of origin e.g. ‘Faransi’ (from France) and ‘Itali’ (from Italy). Some recently introduced cultivars have maintained their original designation e.g. ‘Ferragnès’, ‘Ferraduel’, and ‘Lauranne’.

This study aims at deciphering the evolutionary processes related to human activities, involving exchanges and propagation practices, at different temporal and spatial scales. Available evidences drawn from evolutionary biology, archaeology and history, are integrated to reconstruct the series of events that has led to the continuous diversification of the cultivated almond since its domestication, along with its diffusion dynamics in the Mediterranean Basin. We used a population genetics approach with many individuals representing each cultivar sampled from Lebanon; this approach is well adapted to investigate the genetic diversity of cultivated fruit trees, both within and between cultivars. The intensive sampling of Lebanese cultivars was compared with a large set of cultivated almonds collected *in situ* from different Mediterranean regions. Using this sampling approach we: 1) investigate the genetic differentiation of cultivated almonds in Lebanon in the Mediterranean Basin; 2)



Figure I-1. Geographical distribution of the 92 sampled cultivated almond populations.

Figure I-1 was created in ArcGIS.

characterize the genetic diversity of each identified genetic group; 3) evaluate the gene flow dynamics.

Material and methods

Study sites and samples

The study sample combined materials from Lebanon and from other Mediterranean countries. The Lebanese subsample was obtained from a survey carried out in 2009, 2010 and 2014, searching for both seed and clonally propagated cultivated almond trees, and gathering as many cultivars as possible. Cultivar identification followed farmers' recognition and nomenclature. For common cultivars, sampled orchards were selected based on the availability of 16 to 30 individuals. Rare cultivars were only found in multi-varietal orchards, so that all trees found had to be sampled. By the end of the survey and sampling process, we retained 823 individuals, belonging to 11 cultivars (i.e., 11 vernacular names), collected from 37 orchards across the country (Additional file 1: Table I-S1). Another subsample of 1055 individuals, collected from old orchards in the eastern and western parts of the Mediterranean Basin, was retained from a previous study (Delplancke et al., 2013); neither Tunisian nor Egyptian local almonds were represented in this sample, and local cultivar names could not be recovered as farmers had not been interviewed.

Thus, a total of 1878 individual trees were analyzed from 92 different geographical sites distributed over 12 countries from the east and west of the Mediterranean Basin (Figure I-1).

DNA extraction and microsatellite amplification

Young leaves were collected and dried with silica gel, then maintained at room temperature for DNA extraction. Total genomic DNA was extracted from silica-dried leaves using

Extract-N-Amp™ Plant PCR Kits (SIGMA - Aldrich, St. Louis, MO, USA) abiding by the supplier's instructions.

Individuals were genotyped using 12 previously published nuclear microsatellites shown to be polymorphic in cultivated almond and other *Prunus* species: UDP96-001, UDP96-018, UDP96-003, UDP97-401, UDP98-408, UDP98-409, pchgms1, BPPCT017, BPPCT001, BPPCT007, BPPCT025, BPPCT036 (Cipriani et al., 1999; Testolin et al., 2000; Dirlewanger et al., 2002). We amplified these 12 SSR markers into three multiplexed PCRs, using one of the FAM, HEX or NED fluorophore-labeled primers (PE Applied Biosystems, Warrington, UK). Multiplexed PCRs were carried out with the Extract-N-Amp PCR Ready Mix (SIGMA - Aldrich, St. Louis, MO, USA) in a final volume of 10 µl, containing 5µl of a SIGMA Master Mix 2X, 0.4µl of primer mix at 5µM, 1.6 µl of ultrapure water and 1µl of template dNTPs. 3µl of PCR product was mixed with 15µl of formamide and 0.2µl of Genescan™ 500 LIZ size standard (Applied Biosystems, Foster City, USA), and GeneScan was performed with the ABI PRISM 3130 XL 16 capillary-sequencer (Applied Biosystems, Foster City, USA).

Genotyping and data organization

Allele size was read independently by two investigators using GENEMAPPER 4.0 (Applied Biosystem, Foster City, USA). Genotyping errors were evaluated by checking the reproducibility of migrations using samples replicated on the different plates; in all, 95.1% of tested individuals gave the same allele size (i.e., the error rate was under 0.05). Genotyping failure was checked and the frequency of null alleles for SSR loci was tested using the Expectation Maximization (EM) algorithm (Dempster, Laird and Rubin, 1977) implemented in GENEPOP (Raymond and Rousset, 1995; http://genepop.curtin.edu.au/genepop_op8.html). The same software was used to test each pair of loci for genotypic linkage disequilibrium using the exact test (http://genepop.curtin.edu.au/genepop_op2.html). *P*-values were

estimated using the Markov Chain algorithm with 10000 dememorization steps, 1000 batches and 10000 iterations.

Redundant genotypes were filtered to identify multilocus genotypes (MLGs) using the poppr R package (Kamvar, Tabima and Grünwald, 2014). Identical MLGs were considered as a clone and a network-like representation of private and common clones, among cultivars and sampled regions, was produced using GEPHI 0.7 alpha2 (Bastian, Heymann and Jacomy, 2009).

Data analysis

The analysis methods were chosen to deal with the presume deviation from HWE and LD in the dataset.

Genetic Structure analysis, identification of genetic clusters

The matrix of MLGs was used to perform clustering analysis. We chose to perform the DAPC clustering approach as this method does not rely on any assumptions about population genetics model and could be more efficient than Bayesian model-based approaches to detect complex patterns of genetic differentiation (Jombart et al., 2010). This multivariate analysis seeks linear combinations of the original variables (alleles), which maximize between-group variation and minimize within-group variation. DAPC transforms the data using PCA, and then performs a Discriminant Analysis on the principal components (PC). This prior transformation ensures that variables subjected to the DA are uncorrelated. The identification and description of groups of genetically related MLGs was performed using the Adegenet 2.1.0 R package (Jombart et al., 2017). DAPC analysis was first run without any a priori group assignment. Genetic groups were identified using K-means, a clustering algorithm which finds a given number (K) of groups maximizing the differentiation between groups. To perform the K-means clustering, we retained a subset of nine loci that exhibited frequencies of

null alleles lower than 8%. However, we decided to retain the 12 SSR markers for DAPC analysis to avoid the loss of power of assignation of individuals to different groups. Indeed, Carlsson (2008), using other approaches based on allele frequencies has showed a slight influence on accurately assigning individuals to subpopulations.

The optimal number of groups, K-means was run sequentially with increasing values of K, in a loop of 20 repetitions for clusters ranging from 1 to 20, using the *find.clusters* function. This optimal number of K groups was identified by its lowest Bayesian Information Criterion (BIC) (Schwarz, 1978). We performed a cross-validation to provide an objective optimization procedure for identifying the number of PC axes to retain. Indeed, retaining too many components with respect to the number of individuals can lead to overfitting and instability. The optimization procedure was carried out using the function *xvalDapc* with 100 repetitions proposed by the *adegenet* R package. The function *xvalDapc* performs stratified cross-validation of DAPC using varying numbers of PCs (and keeping the number of discriminant functions fixed). The DAPC analysis was run with a priori assignation based on the retained discriminant functions and the optimal number of groups. Since almond is allogamous and originated from a single gene pool, it is expected that individuals assigned to different groups would have numerous alleles in common; thus, admixed MLGs were identified at a threshold of 0.85 in order to limit misclassifying of not admixed individuals as admixed.

The reassignment test implemented in GENECLASS2 (Piry et al., 2004) was used to verify the success of DAPC, using the Cavalli-Sforza & Edward distance-based method. The 1585 individuals were affected to nine reference populations constructed of individuals assigned to a DAPC group at a threshold ≥ 0.99 ; the level of affectation success was set to 0.05. Probabilities of each individual to belong to reference populations were provided by performing a power analysis by analyzing simulated dataset of 1.000 simulated individuals

($\alpha = 0.01$), using the MCMC and the algorithm of Paetkau et al. (2004). The 12 SSR markers were retained for this analysis.

Genetic diversity and genetic differentiation

The number of different alleles (N_a), allelic richness (Ar) corrected for sample size via bootstrapping, and expected heterozygosity (H_E) were calculated for DAPC groups using the DivBasic function implemented in the DiveRsity R Package (Keenan et al., 2013). Differences among mean diversities between DAPC groups were compared using the Wilcoxon rank sum U-test implemented in R; P-value was adjusted using Bonferroni method.

We used two types of metrics to calculate the population structure: a fixation measure F_{ST} (Weir and Cockerham, 1984) corrected for null alleles, and an allelic differentiation measure D (Jost, 2008). Both types of metrics quantify complementary aspects of population structure. The pairwise unbiased fixation index (F_{ST}^{ENA}) was estimated among DAPC groups using the null alleles adjustment in FreeNA (Chapuis and Estoup, 2007).

The differentiation based on D is used to rank pairs of groups according to their degree of current allelic differentiation. D weighs alleles according to the square of their relative frequencies, so it mainly measures the differentiation of the most common alleles. For each pair of groups, a high value of D indicates that the common alleles are not shared and groups are determined by private alleles, while zero value indicates that the common alleles are shared and have similar frequencies in both groups. However, when the common alleles are shared and some are not, sometimes adding a new low-frequency private allele to a group can slightly reduce the D by reducing the squared relative frequency of more common unshared alleles in a group (Jost et al., 2018).

Since D is sensitive to locus mutation rate, we followed Jost's suggestion (Jost et al., 2018) and selected a subset of six markers that followed a stepwise mutational model according to Delplancke et al. (2012, 2013) and frequencies of null alleles lower than 8%. Therefore, Jost's D between DAPC groups was calculated using the subset of the following six SSR markers (UDP98-408, Pchgms3, BPPCT007, BPPCT017, UDP96-018 and UDP96-003), with 1000 bootstrap and using the diffCalc function as implemented in diveRsity R package (Keenan et al., 2013).

Results

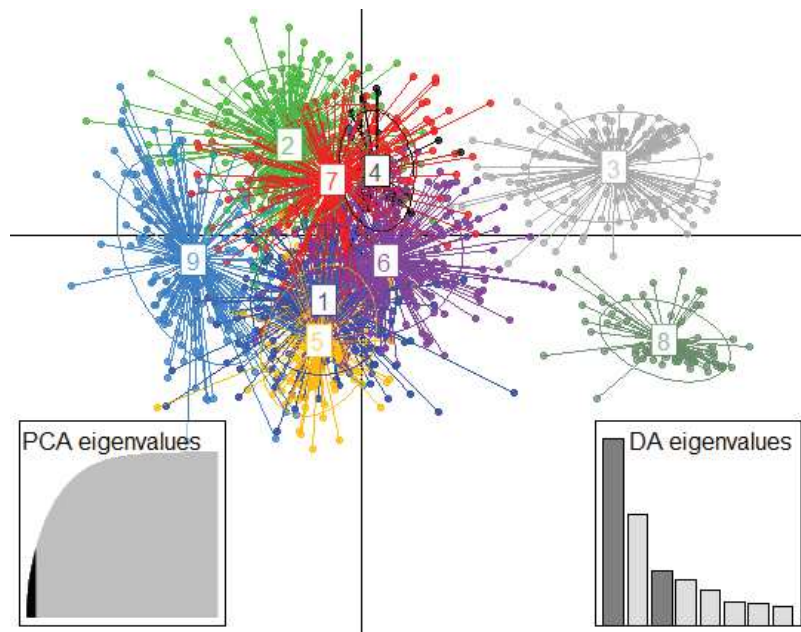
SSR characterization

Summary diversity statistics and null alleles for the 12 loci were calculated for the retained samples (Additional file 1: Table I-S2). Null alleles higher than 8% were significantly present in three loci with an expected frequencies ranging from 15% to 24%. However, no alleles within these loci presented high expected frequencies of pairing with a null allele (data not shown). Significant deviation from Hardy-Weinberg equilibrium was observed for all loci even after Bonferroni correction ($p < 0.004$ at $\alpha = 0.05$) (Additional file 1: Table I-S2). Based on the overall dataset, nine out of 66 locus combinations showed significant linkage disequilibrium after performing Bonferroni correction (Additional file 1: Table I-S3). This genotypic linkage disequilibrium indicates the presence of sub-structuring in the dataset.

Relatedness among samples and clone diffusion

Over the 1878 sampled trees, 380 trees had 87 MLGs in common (clones) among which 74 clones were private to one of the Lebanese cultivars (Additional file 2: Figure I-S1). Four clones were recognized from MLG shared among different Lebanese cultivars. Furthermore, two clones were common to the cultivar Faransi and the French cultivar Ferraduel collected from a Lebanese orchard. One clone corresponded to a MLG shared by an individual

A)



B)

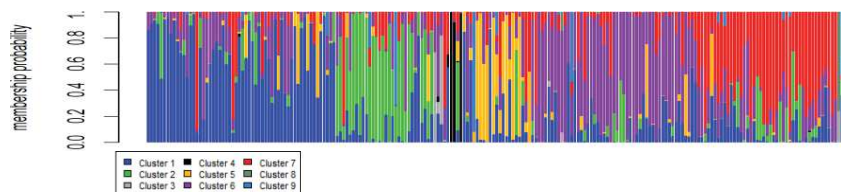


Figure I-2. Discriminant Analysis of Principal Components (DAPC).

A) scatterplot of the final DAPC model, where points are individual MLGs. Since the same alleles were discriminant on the first two axes, we represented the biplot of the first and thirds PCs. The first and third discriminant factors ($PC1 = 37.5\%$ and $PC3 = 10.9\%$ of retained variation) are plotted as the x- and y- axis. DA and PCA Eigenvalues represent the amount of genetic variation captured by the analysis. The bottom right graph illustrates the variation explained by the 19 PCs. B) Structure-like representation of admixed individuals at a cutoff 0.85 resulting from DAPC analysis. Each column corresponds to a MLG and colors correspond to DAPC groups.

collected from Iran and ‘Ferragnès’ and ‘Ferraduel’ individuals collected from Lebanese orchards. Three clones were common to Algeria and Morocco, two clones were common to Algeria and Syria and one clone was common to France and Syria.

DAPC analysis and clustering validation

The K-mean clustering identified $K = 9$ as the optimal number of clusters by retaining 300 axes that represented more than 98% of the total variance. The cross-validation procedure showed that retaining 19 principal component axes give the highest assignment success, representing 41% of the total variance (Additional file: Figure I-S2). The final DAPC analysis was run using the nine clusters obtained by BIC, 19 retained PCs and the eight discriminant axes (Figure I-2). A total of 94.4% of consistency was achieved between assignments prior and posterior to the cross-validation. Individuals from the same country form preferential groupings (Table I-1; Additional file 1: Table I-S4). A total of 19 alleles contributed to the eight PC axes with variable frequencies among the nine DAPC groups (Additional file 1: Table I-S5). Each DAPC group was characterized by a multi-allelic combination (Table I-2). Four alleles contributed significantly to the two main axes (named as loading: allele 160 of mp1, 125 of mp16, 159 of mp1 and 234 of mp7) (Additional file 2: Table I-S5), and differentiated groups 3, 8 and 9 from the others (Figure I- 2A, B).

Group 3 (G3) consisted of MLGs mainly of the Lebanese cultivars (a total of 141 MLGs distributed as follows: 111 ‘Halwani’, 12 ‘Khachabi’, 11 ‘Itali’, 6 ‘Faransi’, and two MLGs from each of ‘Awja’ and ‘Istanbouli’). It also includes a few MLGs collected from Syria (4 MLGs), Algeria (6 MLGs), France (3 MLGs) and Morocco (only one MLG) (Additional file 1: Table I-S4). The assignment effectiveness to this group was 98.8%.

Group 8 (G8) included 70 MLGs sampled from different Lebanese cultivars and only seven MLGs from Syria. All sampled MLGs from cultivar Bandouk Awja and the majority from

Table I-1. The distribution of MLGs sampled from each country over the nine DAPC groups, posterior to cross-validation.

| | G1 | G2 | G3 | G4 | G5 | G6 | G7 | G8 | G9 | Total |
|--------------|------------|------------|------------|-----------|-----------|------------|------------|-----------|-----------|-------|
| Algérie | 1 | 26 | 6 | | | 2 | 11 | | 14 | 60 |
| Maroc | 8 | 152 | 1 | 14 | | 11 | 4 | | 50 | 240 |
| Espagne_Maj | | 21 | | | | 2 | 2 | | 7 | 32 |
| France | 2 | 34 | 3 | 2 | 6 | 8 | 124 | | 30 | 209 |
| Italie_BR | 2 | 2 | | | | 5 | 6 | | 1 | 16 |
| Grece | 6 | | | | 1 | 10 | 4 | | | 21 |
| Turquie | 54 | 8 | | | 60 | 65 | 19 | | 15 | 221 |
| Iran | 4 | | | | 39 | 5 | 10 | | 1 | 59 |
| Turkmenistan | 1 | | | | 3 | 2 | | | 1 | 7 |
| Ouzbekistan | | | | | 4 | | | | 1 | 5 |
| Syrie | 30 | 4 | 4 | | 3 | 24 | 14 | 7 | 9 | 95 |
| Liban | 164 | 26 | 141 | 21 | 3 | 155 | 49 | 70 | 60 | 689 |
| Total | 264 | 250 | 155 | 37 | 118 | 272 | 231 | 77 | 181 | 1585 |

Table I-2. Frequencies of the 19 alleles discriminating among DAPC group.

| Marker | Allele | G1 | G2 | G3 | G4 | G5 | G6 | G7 | G8 | G9 | Contribution on DA |
|--------|--------|------|------|------|------|------|------|------|------|------|--------------------|
| mp16 | 125 | 0.06 | 0.02 | 0.54 | - | 0.03 | 0.08 | 0.02 | 0.79 | 0.03 | PC1 |
| mp51 | 159 | 0.05 | 0.02 | 0.39 | 0.01 | 0.04 | 0.09 | 0.02 | 0.62 | - | PC1 |
| mp51 | 160 | 0.06 | 0.04 | - | - | 0.02 | 0.02 | 0.08 | - | 0.99 | PC1, PC2 |
| mp7 | 230 | 0.29 | 0.16 | 0.89 | 0.26 | 0.22 | 0.67 | 0.17 | 0.97 | 0.30 | PC1, PC8 |
| mp1 | 100 | 0.22 | 0.70 | 0.50 | 0.46 | 0.64 | 0.64 | 0.71 | 0.07 | 0.54 | PC3, PC7 |
| mp7 | 234 | 0.22 | 0.73 | 0.03 | 0.04 | 0.06 | 0.08 | 0.19 | 0.01 | 0.39 | PC3, PC4 |
| mp16 | 105 | 0.24 | 0.18 | 0.09 | 0.95 | 0.40 | 0.16 | 0.49 | 0.03 | 0.24 | PC4 |
| mp17 | 183 | 0.13 | 0.12 | 0.68 | 0.81 | 0.05 | 0.11 | 0.30 | 0.49 | 0.21 | PC4 |
| mp11 | 154 | 0.12 | 0.19 | 0.06 | - | 0.01 | 0.03 | 0.14 | 0.40 | 0.08 | PC5, PC8 |
| mp66 | 138 | 0.03 | 0.09 | 0.12 | 0.53 | 0.02 | 0.07 | 0.07 | 0.53 | 0.08 | PC5 |
| mp66 | 153 | 0.22 | 0.20 | 0.41 | 0.03 | 0.06 | 0.15 | 0.08 | 0.01 | 0.16 | PC5 |
| mp51 | 157 | 0.11 | 0.10 | 0.01 | 0.39 | 0.06 | 0.19 | 0.06 | 0.15 | - | PC6, PC8 |
| mp7 | 238 | 0.08 | 0.01 | - | - | 0.50 | 0.04 | 0.05 | 0.00 | 0.05 | PC6 |
| mp19 | 179 | 0.20 | 0.06 | 0.01 | - | 0.12 | 0.27 | 0.37 | 0.01 | 0.21 | PC7, PC8 |
| mp19 | 177 | 0.35 | 0.25 | 0.02 | - | 0.35 | 0.10 | 0.06 | 0.01 | 0.14 | PC7 |
| mp56 | 137 | 0.19 | 0.29 | 0.42 | 0.50 | 0.07 | 0.09 | 0.19 | 0.01 | 0.23 | PC7 |
| mp8 | 103 | 0.12 | 0.28 | 0.91 | 0.10 | 0.12 | 0.37 | 0.21 | 0.61 | 0.22 | PC7 |
| mp7 | 236 | 0.17 | 0.04 | 0.01 | 0.11 | 0.17 | 0.07 | 0.33 | 0.01 | 0.12 | PC8 |
| mp74 | 156 | 0.03 | 0.04 | 0.45 | - | 0.01 | 0.08 | 0.32 | 0.34 | 0.08 | PC8 |

cultivar Itali were assigned to this group. It contains also MLGs of cultivars Halwani (30 MLGs), Khachabi and Awja (5 MLGs from each cultivar). The assignment effectiveness was 98.7%.

Group 9 (G9) was established with 50 MLGs from Morocco, 30 from France, 15 from Turkey, 14 MLGs from Algeria, seven from Spain, nine from Syria and only one MLG from each of Italy, Iran, Turkmenistan and Uzbekistan. It also includes 60 MLGs sampled from Lebanon: 52 MLGs sampled from different orchards of the cultivar Khachabi, three MLGs from each of 'Awja' and 'Itali' and two MLGs from 'Halwani'. The assignment effectiveness was 99.4%. The results of allelic contribution to the DAPC groups revealed that the allele 160 of locus mp51 was almost exclusive to G9. However, even when we replaced allele 160 by the common allele 159, the G9 group persisted and the number of groups remained equal to $K=9$, only the re-assignment of some individuals to the previous groups was affected (data not shown).

As for the six other DAPC groups, they were differentiated by frequencies of the 15 alleles contributing significantly on the six remaining PC axes (see Table I-2 for multi-alleles combination of each group).

Group 1 (G1) included 164 MLGs of cultivar Khachabi collected mainly from the Bekaa Valley, clustered together with 54 MLGs from Turkey, 30 MLGs from Syria and 6 from Greece. It also included 18 MLGs from other localities: eight from Morocco, four from Iran, two from each of France and Italy and only one MLG from each of Algeria and Turkmenistan. The assignment effectiveness was 93.6%.

Group 2 (G2) included the majority of the samples from Morocco (152 MLGs), Spain (21 MLGs) and Algeria (26 MLGs), associated to 34 MLGs from France and two from Italy. Individuals collected from the eastern parts of the Mediterranean Basin were few and

represented by 26 MLGs from Lebanon (24 and two MLGs from ‘Khachabi’ and ‘Halwani’, respectively), eight MLGs from Turkey and four MLGs from Syria. The assignment effectiveness was 94%.

Group 4 (G4) clustered the 21 MLGs of the cultivar Nahali with 14 MLGs from Morocco and only two MLGs from France. The assignment effectiveness was 94.6%.

Group 5 (G5) clustered together 60 MLGs from Turkey, 39 MLGs from Iran, 3 MLGs from Turkmenistan and 4 MLGs from Uzbekistan. It also includes few MLGs from other countries: 6 MLGs from France, 3 MLGs from each of Syria and Lebanon (‘Khachabi’), and only one MLG from Greece. The assignment effectiveness was 86.4%.

Group 6 (G6) clustered together 155 MLGs of cultivar Khachabi collected mainly from the North, South and Mount Lebanon regions and the cultivar Awja. It also included 24 MLGs from Syria, 65 MLGs from Turkey and 10 MLGs from Greece. Few MLGs from other localities were assigned to this group: eleven from Morocco, eight from France, five from each of Iran and Italy and two from each of Algeria, Spain and Turkmenistan. The assignment effectiveness was 94.8%.

Group 7 (G7) clustered 124 MLGs collected from France together with 11 MLGs from Algeria, 14 MLGs from Syria, 19 MLGs from Turkey, 10 MLGs from Iran, four MLGs from each of Italy and Morocco and only two MLGs from Spain. It also grouped all cultivars originated from France and collected from Lebanese orchards, with 21 MLGs sampled from different ‘Khachabi’ orchards and 7 MLGs from cultivar Halwani. The assignment effectiveness was 90.9%.

A total of 216 admixed MLGs were identified at the threshold of 0.85 (Figure I-2B). The highest admixture level was observed within G1 and G6, followed by G7 and G2 then G5. In contrast, low admixture was observed in the three groups differentiated on the first two PC

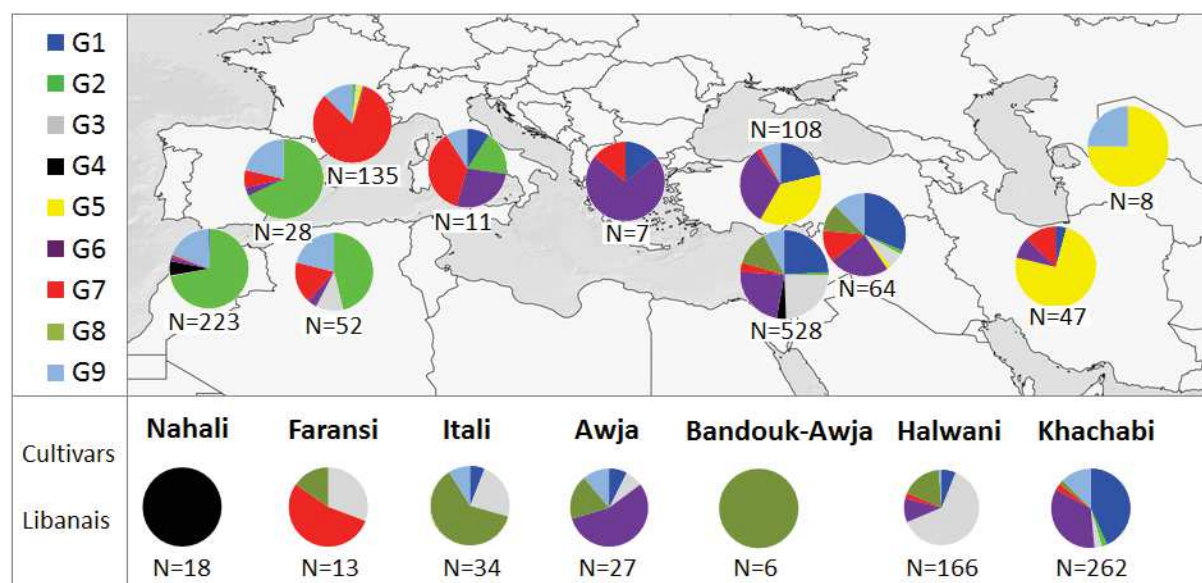


Figure I-3. Geographical distribution of the nine DAPC groups after elimination of admixed and migrant individuals.

Each country is represented in a pie chart on the map, Lebanese cultivars are represented below the line.

Table I-3. Genetic diversity estimations for DAPC groups.

| | N | A | Ar | Ar std | He | He std |
|----|-----|-----|--------------------|--------|-------|--------|
| G1 | 270 | 300 | 16.25 ^a | 3.27 | 0.90 | 0.04 |
| G2 | 242 | 254 | 13.18 ^b | 4.04 | 0.81 | 0.16 |
| G3 | 156 | 152 | 7.35 ^c | 3.45 | 0.59* | 0.22 |
| G4 | 35 | 79 | 5.58 ^d | 1.36 | 0.59* | 0.19 |
| G5 | 109 | 257 | 14.49 ^a | 3.59 | 0.83 | 0.11 |
| G6 | 281 | 314 | 15.80 ^a | 3.93 | 0.84 | 0.14 |
| G7 | 222 | 289 | 14.03 ^a | 3.54 | 0.82 | 0.12 |
| G8 | 74 | 84 | 5.06 ^d | 2.24 | 0.54* | 0.11 |
| G9 | 196 | 247 | 13.68 ^b | 5.27 | 0.79 | 0.26 |

Pairwise comparisons using Wilcoxon rank sum test, values sharing a letter are not significantly different. P-value adjustment method: Bonferroni.

axes (G3, G8 and G9). Admixture was mainly observed among individuals present in the same sampled site or country. For example, MLGs collected from France and attributed to G3, showed admixture with group G7 that includes the majority of the French material. No admixed MLGs were present in the group G4.

The GENECLASS2 analysis validates the DAPC clustering with a reassignment success of (75.3%). Out of the 1585 individuals, 18 individuals were not assigned. A total of 1193 of 1585 individuals to their DAPC groups (at $P > 0.05$); the remaining 388 individuals were reassigned to more than one group (Data not shown).

Distribution of genetic groups and organization of cultivars

The geographical distribution of DAPC groups was represented by the assigned individuals with both DAPC and GENECLASS2 (Figure I-3). The genetic groups were differentiated according to eastern and western Mediterranean regions, except for the group G9 which was distributed across the whole studied region. The DAPC groups were distributed as follows: two western Mediterranean groups (G7, a French group and G2, a Moroccan and Spanish group) and three Eastern Mediterranean groups (G1, a Levant group; G6, a group present in Levant, Greece and Italy; G5, a Greater Persian group). Three additional groups G3, G4 and G8 were composed of cultivars introduced to Lebanon.

Genetic diversity among DAPC groups

Group G1 presents the highest values for allelic richness ($Ar = 16.24$) and expected heterozygosity ($He = 0.90$) (Table I-3). No significant differences for the Ar and He values were identified among G1, G2, G5, G6, G7 and G9 groups (Wilcoxon rank sum test, $P > 0.05$). Allelic richness and heterozygosity were significantly lower in G8, G4 and G3 groups in comparison to the other groups (Wilcoxon rank sum test, $P < 0.05$) (Table I-3). This

Table I-4. Pairwise F_{st}^{ENA} (lower to diagonal) and Jost's D (upper to diagonal) among DAPC groups.

| F _{st} ^{ENA} /D _{jost} | G1 | G2 | G3 | G4 | G5 | G6 | G7 | G8 | G9 |
|---|-------------|------|------|-------------|------|------|------|-------------|-------------|
| G1 | | 0.19 | 0.50 | 0.53 | 0.14 | 0.10 | 0.20 | 0.65 | 0.09 |
| G2 | 0.03 | | 0.44 | 0.49 | 0.33 | 0.20 | 0.19 | 0.62 | 0.08 |
| G3 | 0.10 | 0.12 | | 0.52 | 0.61 | 0.44 | 0.43 | 0.22 | 0.47 |
| G4 | 0.10 | 0.12 | 0.17 | | 0.53 | 0.55 | 0.38 | 0.57 | 0.45 |
| G5 | 0.03 | 0.05 | 0.13 | 0.11 | | 0.22 | 0.26 | 0.68 | 0.25 |
| G6 | 0.02 | 0.04 | 0.07 | 0.11 | 0.03 | | 0.12 | 0.55 | 0.07 |
| G7 | 0.03 | 0.03 | 0.11 | 0.09 | 0.03 | 0.03 | | 0.54 | 0.08 |
| G8 | 0.13 | 0.16 | 0.09 | 0.21 | 0.16 | 0.11 | 0.15 | | 0.60 |
| G9 | 0.04 | 0.04 | 0.13 | 0.12 | 0.06 | 0.04 | 0.04 | 0.17 | |

D values are calculated using 7 SSRs selected for similar mutation rate and low null allele frequencies.

reduction is due to the presence of the clonally propagated cultivars collected from Lebanon in these latter groups.

Genetic differentiation among DAPC groups

Pairwise F_{st}^{ENA} values among DAPC groups ranged from 0.02 (G1 - G6) to 0.21 (G4 - G8) (Table I-4). G1 and G6 have fixed many common alleles, whereas G4 and G8 have fixed different alleles. As compared to G8, G1 was the most differentiated group ($F_{st}^{ENA} = 0.13$), while G3 was the closest group ($F_{st}^{ENA} = 0.09$). Significant differences were detected between G3 and G8, on one hand, and G4, on the other hand. These groups constituted of Lebanese clonally propagated cultivars were also significantly differentiated from the other DAPC groups.

Pairwise D values among DAPC groups ranged from 0.07 (G6 – G9) to 0.68 (G5 – G8) (Table I-4). The G9 shared many alleles with G1, G2 and G7 (91 to 92 % of common alleles were identified within the different pairwise group). As to G8, a high allelic differentiation values was identified in the pairwise comparison with G1 and G2; they were represented by 35 and 38 % of common alleles, respectively. The G8 group had the lowest pairwise D value with the G3, with 78% of common alleles between them ($D = 0.22$). G1 had the lowest pairwise D value with G6, with 90% of common alleles between them ($D = 0.10$).

Discussion

This study investigated the evolutionary dynamics of cultivated almond, which has been shaped by many biological and human-driven factors, with wide spatio-temporal dimensions. The cultivar approach was used to investigate the effect of human drivers such as migrations and cultivation practices on genetic diversity. Ours results showed a differentiation among

ancient almonds sampled from different Mediterranean regions and the recent cultivars introduced to Lebanon.

Our results confirmed the east-west pattern of genetic differentiation across the Mediterranean Basin which was identified by Delplancke et al. (2013) using a Bayesian clustering method (STRUCTURE software) to study the evolutionary history of almond. Different key Mediterranean tree crop species showed east-west patterns of genetic differentiation; e.g. olive (Breton et al., 2009 ; Haouane, 2012 ; Chalak et al., 2015 Diez et al., 2015), fig (Khadari et al., 2005), grapevine (Myles et al., 2011 ; Bacilieri et al., 2013) and date palm (Zehdi-Azouzi et al., 2015). Nevertheless, in contrast to the cultivated almond, the east-west pattern in the aforementioned crops resulted from different domestication events which have occurred in eastern and in western wild populations and which were maintained by early vegetative propagation of domesticated gene pools (Diez et al., 2015; Khadari et al., 2005; Bacilieri et al., 2013). According to Haouane (2012), the east-west gradient of olive genetic structure showed decreasing genetic diversity levels, and the central Mediterranean group had an intermediate level of genetic diversity. Even if introgressions occurred between the cultivated and wild olive present in the western Mediterranean Basin, the main cultivated olive cultivar in Morocco was originated from the eastern domesticated olive. Furthermore, the low genetic diversity in the west reveals the occurrence of a bottleneck during the westward diffusion of eastern cultivated olives (Diez et al., 2015). Such a bottleneck was not supported for the cultivated almond present in the Mediterranean Basin.

Genetic diversity and differentiation of cultivated almond in the Mediterranean Basin

The high level of genetic diversity (A_r and H_e) observed in our samples for the Mediterranean Basin group is consistent with the absence of bottleneck effect, reported along the domestication of this species (Gouta et al., 2010; Elhamzaoui et al., 2012; Delplancke et al.,

2013, Distefano et al., 2013; Fernandez i Marti et al. 2009, 2015). A combination of factors associated to fruit tree species was responsible for a limited bottleneck: long generation time, allogamy and broad geographic distribution (Hamrick and Godt, 1990; Hamrick, Godt and Sherman-Broyles, 1992) along with frequent human-driven local selection and exchange events (Miller and Gross, 2011). This common background to long-lived fruit trees associated to a prolonged cultivation of undomesticated forms prompted the high genetic diversity of ancient eastern and western almond genetic groups. Our results concerning the low genetic differentiation and a high level of common alleles among the Mediterranean Basin groups are explained by their single origin as identified by chloroplastic SSR markers (Delplancke et al., 2013), the geographically diffuse domestication of the cultivated almond in the eastern Mediterranean Basin (Weiss et al., 2006; Willcox et al., 2009) and by the seed propagation of cultivated almond. Furthermore, recurrent hybridization events with sympatric almond wild relatives (Delplancke et al., 2016) may have contributed to increase the genetic diversity level in the east Mediterranean gene pool (G1) since hybridizations were evidenced between the cultivar Khachabi and the sympatric wild relative *P. orientalis* present in the Bekaa Valley region (Delplancke et al., 2012). Moreover, even if limited, gene flow among the eastern groups (e.g. G1 and G6), together with the introduction of genetically differentiated cultivars, have contributed to increase the genetic diversity of the eastern Mediterranean gene pool, mainly in Lebanon, by introducing new alleles.

Differentiation measurements (F_{st}^{ENA} and D_{jost}) indicated that G9 was similarly differentiated from eastern and western Mediterranean Basin groups and much more differentiated from G5 (Greater Persian group). The G9 group had 75% common alleles with the G5, and appeared as secondary gene pool composed by material from the other Mediterranean groups. The presence of such a grouping in G9 revealed complex diffusion which may result from recent admixture among the eastern and western Mediterranean Basin groups. Similarly, an

heterogeneous group identified throughout the areas of olive cultivation in the Mediterranean Basin was originated from recent admixture among distinct groups according to Diez et al. (2015).

The genetic differentiation results highlighted lowest values between G1 and G5 ($F_{st}^{ENA} = 0.03$ and $D_{Jost} = 0.14$), supporting that the most ancient cultivated almond is present in these groups. The comparison of pairwise F_{st}^{ENA} and D_{Jost} values revealed that the western groups were closer to the Levant group G1, than to the Greater Persian group G5. We suggest that the other Mediterranean groups may have originated from the Levant group G1 or were subject to more gene flow with G1.

Our results revealed a reduction of the genetic diversity within the groups of cultivars introduced to Lebanon which were assigned to G3, G4 and G8. This reduction may result from the introduction of small numbers of founder individuals of the new cultivars. The vegetative propagation has then limited the gene flow, and thus maintained a low genetic diversity within introduced cultivars and a high genetic differentiation among them, as shown for G3 and G4 which are mainly constituted of cultivars Halwani and Nahali, respectively. The high genetic differentiation among groups including the introduced cultivars and groups of the Mediterranean region could be explained by the selection of limited numbers of genotypes; this could have produced an important genetic drift and therefore a loss or fixation of different alleles in these new cultivars. This explains the different allelic pattern among the groups constituted of new cultivars and the Mediterranean Basin groups (Table I-2).

The dynamics of the gene flow in the Mediterranean Basin: geography and historical evidences

Our results revealed different patterns of relationship among the cultivars sampled from Lebanon. Individual trees from the local cultivar Khachabi belonged to either of the eastern

Mediterranean Basin groups G1 or G6, associated to individuals sampled from Syria, Turkey and Greece; this result extended the geographical distribution of the group C3 (G6 in our study) reported by Delplancke et al (2013) to Lebanon. As to the introduced cultivars our results showed that they are not originating from the studied Mediterranean gene pool, except for the recently introduced French cultivars (Lauranne, Ferragnèse and Ferraduel) which regrouped with the French group (G7). Human-driven diffusions of cultivated almond carried out at different times have favored the presence of differentiated groups in the same region.

Ancient human-mediated diffusions in the eastern Mediterranean Basin and local selections may explain the existence of the G1 and G6 groups in the same eastern geographical region, both including individuals from the cultivar Khachabi. The G1 group was maintained by continuous exchanges among the Bekaa Valley, Syria and southern Turkey, due to the geographical proximity and accessibility to the Bekaa Valley from its north and central eastern sides. Since Phoenician, Greek and Roman times, the Lebanese coastal regions had extensive trading exchanges with different Mediterranean regions; this could have favored exchange of almond among these regions. Furthermore, until the 1914, the inland zone and the coastal Mediterranean zone were under different political regimes: the inland zone was geopolitically assigned to Syria, while the coastal Mediterranean zone was mainly under the Lebanese government. We assume therefore, that G1 represented the original almond diversified in the eastern Mediterranean; a part of this group was introduced to Greece where it was sexually propagated and differentiated into G6. Some of the G6 individuals were then reintroduced to the Levant and cultivated in the Lebanese coastal regions. Our scenario corresponds to local almond differentiation due to the presence of geographical and political barriers which limited the diffusion of almond between the Bekaa Valley and the Lebanese coastal regions. Such a differentiation pattern into two groups, as result of different exchange networks, was also observed for almond in Tunis and Spain. According to Gouta et al.(2016),

the northern Tunisian almonds were differentiated from the central and southern gene pool due to the exchange of genotypes between both Mediterranean shores; while the cultivated almond in the center and south of Tunisia represent the North African gene pool established with early introductions of almond into Africa without experiencing exchanges with the Mediterranean gene pool. Similar pattern of differentiation could also be supported for Spanish almond; our results showed that the almond sampled from Majorca island was grouped with Moroccan almond, however, Halász et al. (2016) showed that the continental Spanish almond was regrouped with the French ones. Indeed, the exchanges of almond among different Mediterranean regions have occurred at different periods. However, it was intensified among the different Mediterranean regions (e.g. France, Italy, Spain and Tunisia) during the last century in the G.R.E.M.P.A. framework (Vargas, 2014) which explains the reduction of the genetic differentiation among the Mediterranean regions and the groups obtained in our study.

Recent human-mediated diffusions are evidenced by common clones identified among different eastern and western regions; this highlights the importance of recent reintroduction events from western into eastern regions. For instance, the introduction of French cultivars backward to the eastern Mediterranean was evidenced by the presence of ‘Ferragnès’, ‘Ferraduel’ and ‘Lauranne’ in Lebanese commercial orchards. Migrations from the western group G7 to the eastern region were in line with the recent eastward diffusion which has been initiated in the last century with the introduction of modern cultivars (Grasselly, 1976). Recently introduced (by 1990th), the vegetative propagation of the French cultivars has prevented, up to now, gene flow among them and older cultivars growing in Lebanon. French cultivars had conserved their own name or the name of their country of origin (Faransi) and were gathered into the G7 group with individuals sampled from France. Our results highlighted also the recent backward diffusion of cultivated almond from the Mediterranean

Basin to Asia; the introduction of French cultivars into Iran was evidenced by the presence of a common clone among MLGs of the cultivars Ferragnès and Ferraduel and one MLG sampled from Iran. Furthermore, the distribution of particular clones across Syria, Morocco and other countries highlighted a long distance diffusion of almond, even if the origin of these clones could not be identified.

The pattern of contemporary migrations presented in this study is only a snapshot of complex and ongoing interactions influenced by different actors. The intermediate geographical position of some regions has favored the co-existence of genetically differentiated groups resulting from the inter-regional exchanges. Turkey and Italy hold geographical differentiated groups in consistency with their intermediate geographic position. Historical diffusion along the Silk Road going through Turkey is considered a historical trade corridor for agricultural goods between Asia and Mediterranean regions. Thus, the presence of both eastern Mediterranean groups (G1 and G6) and the Greater Persian group (G5) in Turkey may be due to historical almond diffusion from nearby regions. The presence of G7 genotypes in Turkey is the likely result of the introduction of French selected almond cultivars during the last century; this would be consistent with the finding of Grasselly (1976) who stated that French material was diffused into different regions of the Mediterranean Basin. As to Italy, ancient traders may have introduced individuals from the group G6 while the western Mediterranean Basin group G7 may have developed locally.

Sexual propagation has enhanced introgression among different groups and reduced the genetic differentiation among them. Introgressions by pollen-mediated gene flow among the western groups (G7 and G2) and the Greater Persian group (G5) with both eastern Mediterranean Basin groups (G1 and G6) were identified in orchards where these groups co-existed. Admixture among different almond genetic groups of local and introduces cultivars was evidenced in Lebanon among ‘Khachabi’ and ‘Halwani’ (Hamadeh et al. 2018), it was

also indicated in Sardinia and Iran (Rigoldi et al., 2015; Fathi et al., 2008). Furthermore, this study revealed local introgression by pollen-mediated gene flow among different introduced cultivars, without excluding some farmer's mislabeling. Indeed, the cultivar Halwani was mainly assigned to a genetically differentiated group (G3). Nonetheless, a substantial part of the individuals from 'Halwani' did not belong to G3, constituting the G8 group along with individuals from other cultivars (e.g. Awja, Bandouk Awja, Itali and Faransi). According to farmers, the cultivar Bandouk Awja resulted of local hybridizations among the cultivar Awja and probably the cultivar Halwani; this could explain the assignation of these three cultivars to a single group (G8). The genetic proximity of the three aforementioned cultivars with the cultivars Itali and Faransi supports a local introgression by pollen-mediated gene flow.

Finally, pollen-mediated gene flows have been reduced by the shift to vegetative propagation for introduced cultivars, limiting the introgression among G3, G4 and G8 on one hand, and, on the other hand, between them and Mediterranean Basin groups G1, G2, G6, and G7. The effect of the history of diffusion and the vegetative propagation practices could be attested by the differentiation of the cultivar Nahali into the G4 group; a strong founding event followed by restricted geographical diffusion and vegetative propagation have limited gene flows with other cultivars. According to farmers, three seeds were introduced from America, early in the last century by a religious named Nahli, planted and then propagated vegetatively in the Batroun region (North Lebanon).

Conclusion

Almond is a relevant model for deciphering the evolutionary processes related to human activities at different temporal and spatial scales. Archeological, historical and genetic evidences have previously demonstrated that almond domestication involved multiple spatial and temporal steps.

The cultivar approach adopted in this study allowed identifying the recent directional gene flows among regions taking into account the human driven dynamics. Gene flow mainly occurred among geographically close eastern groups. The introduction of foreign cultivars from different origins, and the presence of local traditional orchards, has maintained a substantial genetic diversity in Lebanon. Such introgression has been an important component enhancing the genetic diversity and limiting genetic structuration of cultivated almond. Finally, using neutral markers, this study allowed a better understanding of the historical dynamics of almond genetic diversity. A future genomic approach may help to identify genes subject to selection along the domestication process and ascertain proposed scenarios.

References

- Albala, K. (2009) 'Almonds Along the Silk Road: The Exchange and Adaptation of Ideas from West to East', *Petits Propos Culinaire*, pp. 19–34.
- Bacilieri, R. *et al.* (2013) 'Genetic structure in cultivated grapevines is linked to geography and human selection', *BMC Plant Biology*, 13(1). doi: 10.1186/1471-2229-13-25.
- Bastian, M., Heymann, S. and Jacomy, M. (2009) 'Gephi: An Open Source Software for Exploring and Manipulating Networks', *Third International AAAI Conference on Weblogs and Social Media*, pp. 361–362. doi: 10.1136/qshc.2004.010033.
- Breton, C. *et al.* (2009) 'The origins of the domestication of the olive tree', *Comptes Rendus - Biologies*, 332(12), pp. 1059–1064. doi: 10.1016/j.crv.2009.08.001.
- Browicz, K. and Zohary, D. (1996) 'The genus *Amygdalus* L. (Rosaceae): Species relationships, distribution and evolution under domestication', *Genetic Resources and Crop Evolution*, 43(3), pp. 229–247. doi: 10.1007/BF00123275.
- Carlsson, J. (2008) 'Effects of microsatellite null alleles on assignment testing', *Journal of Heredity*, 99(6), pp. 616–623. doi: 10.1093/jhered/esn048.
- Certal, A. C. *et al.* (2002) 'Structural and molecular analysis of self-incompatibility in almond (*Prunus dulcis*)', *Sexual Plant Reproduction*, 15(1), pp. 13–20. doi: 10.1007/s00497-002-0138-4.
- Chalak, L. *et al.* (2006) 'PRELIMINARY CHARACTERIZATION OF CULTIVATED ALMONDS (*PRUNUS DULCIS* L.) IN LEBANON BY MORPHOLOGICAL TRAITS AND MICROSATELLITE MARKERS', *Biologia Tunisie*, 4.
- Chalak, L. *et al.* (2015) 'Extent of the genetic diversity in Lebanese olive (*Olea europaea* L.) trees: a mixture of an ancient germplasm with recently introduced varieties', *Genetic Resources and Crop Evolution*, 62(4), pp. 621–633. doi: 10.1007/s10722-014-0187-1.
- Chapuis, M. P. and Estoup, A. (2007) 'Microsatellite null alleles and estimation of population differentiation', *Molecular Biology and Evolution*, 24(3), pp. 621–631. doi: 10.1093/molbev/msl191.

- Cipriani, G. *et al.* (1999) 'AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L) Batsch]: isolation, characterisation and cross-species amplification in *Prunus*', *Theoretical and Applied Genetics*. Springer, 99(1–2), pp. 65–72.
- Delplancke, M. *et al.* (2012) 'Gene flow among wild and domesticated almond species: Insights from chloroplast and nuclear markers', *Evolutionary Applications*, 5(4), pp. 317–329. doi: 10.1111/j.1752-4571.2011.00223.x.
- Delplancke, M. *et al.* (2013) 'Evolutionary history of almond tree domestication in the Mediterranean basin', *Molecular Ecology*. Wiley/Blackwell (10.1111), 22(4), pp. 1092–1104. doi: 10.1111/mec.12129.
- Delplancke, M. *et al.* (2016) 'Combining conservative and variable markers to infer the evolutionary history of *Prunus* subgen. *Amygdalus* s.l. under domestication', *Genetic Resources and Crop Evolution*, 63(2), pp. 221–234. doi: 10.1007/s10722-015-0242-6.
- Dempster, A. P., Laird, N. M. and D.B. Rubin (1977) *Maximum Likelihood from Incomplete Data via the EM Algorithm*, *Journal of the Royal Statistical Society. Series B (Methodological)*. doi: 10.2307/2984875.
- Denisov, V. P. (1988) 'Almond genetic resources in the USSR and their use in production and breeding', *Acta Hort.*, 224, pp. 299–306.
- Diez, C. M. *et al.* (2015) 'Olive domestication and diversification in the Mediterranean Basin', *New Phytologist*, 206(1), pp. 436–447. doi: 10.1111/nph.13181.
- Dirlewanger, E. *et al.* (2002) 'Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.)', *Theoretical and Applied Genetics*, 105(1), pp. 127–138. doi: 10.1007/s00122-002-0867-7.
- Distefano, G. *et al.* (2013) 'Genetic diversity and relationships among Italian and foreign almond germplasm as revealed by microsatellite markers', *Scientia Horticulturae*, 162, pp. 305–312. doi: 10.1016/j.scienta.2013.08.030.
- Donghuai, S. *et al.* (1998) 'Magnetostratigraphy and palaeoclimatic significance of Late Tertiary aeolian sequences in the Chinese Loess Plateau', *Geophysical Journal International*. Oxford University Press, 134(1), pp. 207–212. doi: 10.1046/j.1365-246x.1998.00553.x.
- Elhamzaoui, A. *et al.* (2012) 'Assessment of Genetic Diversity of Moroccan Cultivated Almond (*Prunus dulcis* Mill . DA Webb) in Its Area of Extreme Diffusion , Using Nuclear Microsatellites', *American Journal of Plant Sciences*, 3, pp. 1294–1303.
- Fathi, A., Ghareyazi, B., Haghazari, A., Ghaffari, M.R., Pirseyedi, S.M., Kadkhodaei, S. and Naghavi, M.R. (2008) 'Assessment of the genetic diversity of almond (*Prunus dulcis*) using microsatellite markers and morphological traits', *IRANIAN JOURNAL of BIOTECHNOLOGY*, 6(2), pp. 98–106.
- Felipe, A. J. and Socias I Company, R. (1977) 'Un amandier sauvage, probablement *A. webbii*, non encore mentionné en Espagne', *3d Coll. GREMP, CHEAM, Bari, Italy*, pp. 78–79.
- Fernandez i Marti, A. *et al.* (2009) 'Genetic Diversity in Spanish and Foreign Almond Germplasm Assessed by Molecular Characterization with Simple Sequence Repeats', *Journal of American Society Horticulture Science*, 134(5), pp. 535–542. Available at: <http://journal.ashspublications.org/content/134/5/535.short>.
- Godini, A. (1979) 'Ipotesi sulla comparsa dell'autocompatibilità nel mandorlo', *Riv. Sci. Tecn. Agr.*
- Godini, A. (2000) 'About the possible relationships between *Amygdalus webbii* Spach and *Amygdalus*

communis L.', *NUCIS Newsletter*.

Gouta, H. *et al.* (2010) 'Assessment of genetic diversity and relatedness among Tunisian almond germplasm using SSR markers', *Hereditas*. Wiley Online Library, 147(6), pp. 283–292.

Gouta, H. *et al.* (2016) 'Contribution in tracing the origin and evolutionary history of almond around the Mediterranean basin', *Options Méditerranéennes*, Série A. S(119), pp. 93–97. Available at: <http://om.ciheam.org/om/pdf/a119/00007371.pdf> (Accessed: 24 June 2018).

Gradziel, T. M. (2011) 'Origin and Dissemination of Almond', in *Horticultural Reviews*. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 23–81. doi: 10.1002/9780470872376.ch2.

Gradziel, T. M. (2017) 'History of cultivation.', in *Almonds: botany, production and uses*. Wallingford: CABI, pp. 43–69. doi: 10.1079/9781780643540.0043.

Grasselly, C. (1976) 'Les espèces sauvages d'amandiers', *Ciheam-Options Méditerranées*, 43, pp. 28–43.

Grasselly, C. and Crossa-Raynaud, P. (1980) *L'amandier*. Maisonneuv.

Halász J *et al.* (2016) 'Genetic characterization of almond (*Prunus dulcis*) cultivars and natural resources', *Option Méditerranéennes*, 119, pp. 49–53. Available at: <http://om.ciheam.org/article.php?IDPDF=00007362http://www.ciheam.org/> (Accessed: 10 July 2018).

Hamadeh, B. *et al.* (2018) 'Evolution of almond genetic diversity and farmer practices in Lebanon: impacts of the diffusion of a graft-propagated cultivar in a traditional system based on seed-propagation', *BMC Plant Biology*. BioMed Central, 18(1), p. 155. doi: 10.1186/s12870-018-1372-8.

Hamrick, J. L. and Godt, M. J. W. (1990) 'Allozyme diversity in plant species.', in Brown, A. H. D.; Clegg, M. T.; Kahler, A. L.; Weir, B. S. (ed.) *Plant population genetics, breeding, and genetic resources*. Sinauer Associates Inc., pp. 43–63. Available at: <https://www.cabdirect.org/cabdirect/abstract/19901612624>

Hamrick, J. L., Godt, M. J. W. and Sherman-Broyles, S. L. (1992) 'Factors influencing levels of genetic diversity in woody plant species', *New Forests*. Kluwer Academic Publishers, 6(1–4), pp. 95–124. doi: 10.1007/BF00120641.

Haouane, H. (2012) *Origines , domestication et diversification variétale chez l ' olivier (Olea europaea L .) à l ' ouest de la Méditerranée*.

Jombart, A. T. *et al.* (2017) 'R-Package “ adegenet ”'.

Jombart, T. *et al.* (2010) 'Discriminant analysis of principal components: a new method for the analysis of genetically structured populations', *BMC Genetics*, 11(1), p. 94. doi: 10.1186/1471-2156-11-94.

Jost, L. (2008) 'GST and its relatives do not measure differentiation', *Molecular Ecology*, 17(18), pp. 4015–4026. doi: 10.1111/j.1365-294X.2008.03887.x.

Jost, L. *et al.* (2018) 'Differentiation measures for conservation genetics', *Evolutionary Applications*, (December 2017), pp. 1–10. doi: 10.1111/eva.12590.

Kalinowski, S. T. and Taper, M. L. (2006) 'Maximum likelihood estimation of the frequency of null alleles at microsatellite loci', *Conservation Genetics*, 7(6), pp. 991–995. doi: 10.1007/s10592-006-9134-9.

Kamvar, Z. N., Tabima, J. F. and Grünwald, N. J. (2014) '*Poppr*: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction', *PeerJ*, 2, p. e281. doi:

10.7717/peerj.281.

Keenan, K. *et al.* (2013) 'DiveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors', *Methods in Ecology and Evolution*, 4(8), pp. 782–788. doi: 10.1111/2041-210X.12067.

Khadari, B. *et al.* (2005) 'Contrasted genetic diversity and differentiation among Mediterranean populations of *Ficus carica* L.: A study using mtDNA RFLP', *Genetic Resources and Crop Evolution*, 52(1), pp. 97–109. doi: 10.1007/s10722-005-0290-4.

Martinoli, D. and Jacomet, S. (2004) 'Identifying endocarp remains and exploring their use at Epipalaeolithic Öküzini in southwest Anatolia, Turkey', *Vegetation History and Archaeobotany*, 13(1), pp. 45–54. doi: 10.1007/s00334-003-0029-0.

Miller, A. J. and Gross, B. L. (2011) 'From forest to field: Perennial fruit crop domestication', *American Journal of Botany*, 98(9), pp. 1389–1414. doi: 10.3732/ajb.1000522.

Myles, S. *et al.* (2011) 'Genetic structure and domestication history of the grape', *National Acad Sciences*, 108(9), pp. 3530–3535. Available at: <http://www.pnas.org/content/108/9/3530.short> (Accessed: 24 June 2018).

Piry, S. *et al.* (2004) 'GENECLASS2: A software for genetic assignment and first-generation migrant detection', *Journal of Heredity*, 95(6), pp. 536–539. doi: 10.1093/jhered/esh074.

Raymond, M. and Rousset, F. (1995) 'GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism', *Journal of Heredity*. doi: 10.1093/oxfordjournals.jhered.a111573.

Rigoldi, M. P. *et al.* (2015) 'Genetic diversity in two Italian almond collections', *Electronic Journal of Biotechnology*. Elsevier B.V., 18(1), pp. 40–45. doi: 10.1016/j.ejbt.2014.11.006.

Schwarz, G. (1978) 'Estimating the Dimension of a Model', *The Annals of Statistics*, 6(2), pp. 461–464. doi: 10.1214/aos/1176344136.

Socias, R. and Felipe, A. J. (1988) 'SELF-COMPATIBILITY IN ALMOND: TRANSMISSION AND RECENT ADVANCES IN BREEDING', *Acta Horticulturae*, (224), pp. 307–318. doi: 10.17660/ActaHortic.1988.224.38.

Testolin, R. *et al.* (2000) 'Microsatellite DNA in peach (*Prunus persica* L. Batsch) and its use in fingerprinting and testing the genetic origin of cultivars.', *Genome*, 43(3), pp. 512–520. doi: 10.1139/gen-43-3-512.

Vargas, F. J. (2014) 'INFLUENCE OF RESEARCH ON MODERNIZATION OF ALMOND AND PISTACHIO CULTURE IN MEDITERRANEAN AREA DURING THE LAST FORTY YEARS AND SOME CHALLENGES FOR THE FUTURE', *Acta Horticulturae*, (1028), pp. 19–26. doi: 10.17660/ActaHortic.2014.1028.1.

Vavilov, N. (1930) Wild progenitors of the fruit trees of Turkistan and the Caucasus and the problem of the origin of fruit trees, Report. and Proc. 9th Int. Hort. Congr. Available at: <https://www.cabdirect.org/cabdirect/abstract/19311600466>

Watkins, R. (1976) 'Cherry, plum, peach, apricot and almond: *Prunus* spp.(Rosaceae)', in *Evolution of Crop Plants*.

Weir, B. S. and Cockerham, C. C. (1984) 'Estimating F-Statistics for the Analysis of Population Structure', *Evolution*. doi: 10.2307/2408641.

Weiss, E., Kislev, M. E. and Hartmann, A. (2006) 'Autonomous cultivation before domestication', *Science*, 312, pp. 1608–1610. doi: 10.1126/science.1127235.

- Willcox, G., Buxo, R. and Herveux, L. (2009) 'Late Pleistocene and early Holocene climate and the beginnings of cultivation in northern Syria', *The Holocene*. SAGE PublicationsSage UK: London, England, 19(1), pp. 151–158. doi: 10.1177/0959683608098961.
- Willcox, G., Fornite, S. and Herveux, L. (2007) 'Early Holocene cultivation before domestication in northern Syria', *Veghet Hist Archeobot*, 17, pp. 313–325. doi: 10.1007/s00334-007-0121-y.
- Yazbek, M. (2010) 'Systematics Of Prunus Subgenus Amygdalus Monograph And Phylogeny'. Available at: <https://ecommons.cornell.edu/handle/1813/14843> (Accessed: 5 July 2018).
- Yazbek, M. and Oh, S. H. (2013) 'Peaches and almonds: Phylogeny of Prunus subg. Amygdalus (Rosaceae) based on DNA sequences and morphology', *Plant Systematics and Evolution*, 299(8), pp. 1403–1418. doi: 10.1007/s00606-013-0802-1.
- Zeder, M. A. *et al.* (2006) 'Documenting domestication: The intersection of genetics and archaeology', *Trends in Genetics*, 22(3), pp. 139–155. doi: 10.1016/j.tig.2006.01.007.
- Zehdi-Azouzi, S. *et al.* (2015) 'Genetic structure of the date palm (*Phoenix dactylifera*) in the Old World reveals a strong differentiation between eastern and western populations', *Annals of Botany*, 116(1), pp. 101–112. doi: 10.1093/aob/mcv068.
- Zeinalabedini, M. *et al.* (2010) 'The origin and dissemination of the cultivated almond as determined by nuclear and chloroplast SSR marker analysis', *Scientia Horticulturae*, 125, pp. 593–601. doi: 10.1016/j.scienta.2010.05.007.
- Zohary, D. (2004) 'Unconscious selection and the evolution of domesticated plants', *Economic Botany*, 58(1), pp. 5–10. doi: 10.1663/0013-0001(2004)058[0005:USATEO]2.0.CO;2.
- Zohary, D. and Hopf, M. (1993) 'Domestication of Plants in the Old World, 2nd edit'.

Additional file 1

Table I-S1 Sampling effort.

| Population name | Country | Cultivar name | Code | NDD | EDD | N |
|---------------------|---------|---------------|--------|---------|---------|----|
| Bire | Lebanon | Halwani | HB_Bi | 34.5872 | 36.2391 | 23 |
| Ferzol | Lebanon | Halwani | HB_Fe | 33.8884 | 35.9396 | 33 |
| Nabi-Ayla | Lebanon | Halwani | HB_NA | 33.8819 | 35.9569 | 30 |
| Zeghrine | Lebanon | Halwani | HB_Ze | 34.4261 | 36.3528 | 16 |
| Amchit | Lebanon | Halwani | HM_Am | 34.1475 | 35.6464 | 26 |
| Borjein | Lebanon | Halwani | HM_Bo | 33.6562 | 35.4860 | 12 |
| Maasriti | Lebanon | Halwani | HM_Ma | 33.7483 | 35.6347 | 30 |
| Jeita | Lebanon | Halwani | HM_Je | 33.9533 | 35.6478 | 30 |
| Rejmeh | Lebanon | Halwani | HM_Rj | 33.4714 | 35.3665 | 12 |
| Bchannine | Lebanon | Halwani | HN_Bc | 34.3483 | 35.8859 | 32 |
| Btaaboura | Lebanon | Halwani | HN_Bt | 34.2742 | 35.7622 | 20 |
| Deir-Qanoun-el-Nahr | Lebanon | Halwani | HS_DQ | 33.3000 | 35.3136 | 23 |
| Tanbourite | Lebanon | Halwani | HS-Ta | 33.5161 | 35.4153 | 15 |
| Deir Amar P1 | Lebanon | Awja | AN_DA | 34.2758 | 35.5338 | 2 |
| Kfarhata | Lebanon | Awja | AN-Kf | 34.1699 | 35.4467 | 6 |
| Rachaya | Lebanon | Awja | AB-Ra | 33.511 | 35.8531 | 1 |
| Sfarai | Lebanon | Awja | AS-Sf | 33.5777 | 35.5048 | 2 |
| Zouq Mkayel p1 | Lebanon | Awja | AM-Zo | 33.5754 | 35.3743 | 1 |
| Wadi el Rihan | Lebanon | Awja | AN-WR | 34.3073 | 36.0112 | 16 |
| Assia | Lebanon | Awja | AN-As | 34.2219 | 35.7521 | 1 |
| Bourjein | Lebanon | Awja | AM-Bo | 33.6562 | 35.4860 | 1 |
| Deir Qanoun en Nahr | Lebanon | Awja | AS-DQ | 33.3045 | 35.3246 | 3 |
| Kfarhatna | Lebanon | Awja | AN-Kn | 34.2374 | 35.7558 | 1 |
| Mays al Jabal P2 | Lebanon | Awja | AS-MJ2 | 33.1830 | 35.5309 | 3 |
| Mays al Jabal P3 | Lebanon | Awja | AS-MJ3 | 33.1854 | 35.5277 | 2 |
| El Rejmeh | Lebanon | Awja | AM-Rj | 33.472 | 35.3668 | 2 |
| Halat | Lebanon | Awja | AM-Ha | 34.0842 | 35.6542 | 3 |
| Halat | Lebanon | Bandouk-Awja | BM_Ha | 34.0842 | 35.6542 | 30 |
| Haouch-el-Oumara | Lebanon | Ferraduel_L | FB_D | 33.8192 | 35.9327 | 22 |
| Haouch-el-Oumara | Lebanon | Ferragnès_L | FB_G | 33.8192 | 35.9327 | 29 |
| Zahlé | Lebanon | Lauranne_L | FB_L | 33.8555 | 35.8719 | 26 |
| Wadi el Rihan | Lebanon | Faransi | FN_WR | 34.3073 | 36.0112 | 27 |
| Ijdabra | Lebanon | Nahali | NN_Ij | 34.1439 | 35.4146 | 20 |
| Kfarhatna | Lebanon | Nahali | NN_Kf | 34.2370 | 35.7558 | 15 |
| Bednayel | Lebanon | Itali | IN_Be | 34.1723 | 35.4346 | 19 |
| Wadi el Rihan | Lebanon | Itali | IN_WR | 34.3073 | 36.0112 | 24 |
| Wadi el Rihan | Lebanon | Itali | IN_Ba | 34.3073 | 36.0112 | 9 |

| | | | | | | |
|---------------|---------|-----------|-----------|---------|---------|----|
| Amchit | Lebanon | Istanbuli | Istanbuli | 34.0834 | 35.3908 | 2 |
| Bakkifa | Lebanon | Khachabi | KB_Ba | 33.4933 | 35.8192 | 26 |
| Bire | Lebanon | Khachabi | KB_Bi | 33.5842 | 35.8197 | 20 |
| El-Mhaidseh | Lebanon | Khachabi | KB_EM | 33.5569 | 35.8117 | 30 |
| Irsal | Lebanon | Khachabi | KB_Ir1 | 34.1721 | 36.4079 | 18 |
| Irsal | Lebanon | Khachabi | KB_Ir2 | 34.1886 | 36.3923 | 15 |
| Kamed-El-Lawz | Lebanon | Khachabi | KB_KL | 33.6203 | 35.8214 | 30 |
| Kousaya | Lebanon | Khachabi | KB_Ko | 33.86 | 36.06 | 5 |
| Labweh | Lebanon | Khachabi | KB_La | 34.17 | 36.30 | 25 |
| Lucy | Lebanon | Khachabi | KB_Lu | 33.6453 | 35.8383 | 28 |
| Shaat | Lebanon | Khachabi | KB_Sh | 34.1422 | 36.2311 | 17 |
| Mayrouba | Lebanon | Khachabi | KM_Ma | 34.0105 | 35.7668 | 14 |
| Assia | Lebanon | Khachabi | KN_As | 34.2189 | 35.7856 | 15 |
| El-Qalamoun | Lebanon | Khachabi | KN_EQ | 34.3872 | 35.7864 | 16 |
| Michmich | Lebanon | Khachabi | KN_Mi | 34.5439 | 36.1761 | 12 |
| Baraachit | Lebanon | Khachabi | KS_Ba | 33.1761 | 35.4433 | 30 |
| Blida | Lebanon | Khachabi | KS_Bl | 33.1400 | 35.5147 | 29 |
| Chebaa | Lebanon | Khachabi | KS_Ch | 33.3475 | 35.7492 | 23 |
| Mais-El-Jabal | Lebanon | Khachabi | KS_MJ | 33.1686 | 35.5242 | 18 |
| Rmeich | Lebanon | Khachabi | KS_Rm | 33.0792 | 35.4564 | 6 |
| Tanbourite | Lebanon | Khachabi | KS_Ta | 33.5161 | 35.4153 | 30 |

Table I-S2 Statistics associated with microsatellite loci.

| Locus name | UD | | | | | | | | | | | |
|----------------|------------------------------|--------------------------------|--------------------------------|--------------------------------|------------------------------|------------------------------|-------------------------------|-------------------------------|-------------------------------|--------------------------------|-------------------------------|--------------------------------|
| | P96 - 001 ^a | UDP 98- 409 ^b | UDP 97- 401 ^a | UDP 98- 408 ^b | Pch gms 1 ^c | Pch gms 3 ^c | BPP CT0 01 ^d | BPP CT0 07 ^d | BPP CT0 17 ^d | UDP 96- 018 ^a | BPP CT0 25 ^d | UDP 96- 003 ^b |
| Multiple x | 1b | 2 | 2 | 1b | 1b | 1b | 3 | 3 | 2 | 2 | 1b | 2 |
| Primer code | mp1 | mp11 | mp14 | mp16 | mp1 7 | mp1 9 | mp5 1 | mp5 6 | mp6 6 | mp7 | mp7 4 | mp8 |
| NA | 1/15 | 10/15 | 9/15 | 1/15 | 3/15 | 0/15 | 13/1 5 | 1/15 | 1/15 | 1/15 | 1/15 | 0/15 |
| F (Null) | 0.03 | 0.18 | 0.15 | 0.07 | 0.07 | 0.04 | 0.24 | 0.04 | 0.05 | 0.09 | 0.07 | 0.05 |
| Ho | 0.65 | 0.58 | 0.66 | 0.75 | 0.76 | 0.81 | 0.47 | 0.83 | 0.81 | 0.58 | 0.80 | 0.75 |
| He | 0.67 | 0.93 | 0.95 | 0.88 | 0.90 | 0.89 | 0.92 | 0.89 | 0.88 | 0.75 | 0.92 | 0.86 |
| Fis | 0.04 | 0.37 | 0.30 | 0.15 | 0.16 | 0.09 | 0.49 | 0.07 | 0.09 | 0.22 | 0.13 | 0.13 |

NA: presence of null alleles in the retained 15 countries and cultivars. f (Null): frequency of null alleles estimated based on Maximum likelihood EM algorithm (Dempster, Laird and Rubin, 1977). Ho: observed heterozygosity. He: unbiased expected heterozygosity. Fis: inbreeding index based on *Weir & Cockerham* (1984) formula. All loci had significant HW disequilibrium even after Bonferroni correction ($P < 0.004$).

Table I-S3 Estimation of the pairwise LD frequencies calculated among the 12 SSR markers used in this study. Correlation between a pair of loci calculated using the linkage disequilibrium formula and 1000 permutations. P-value = 0, Loci are in complete linkage equilibrium. P-value = 1, Loci are in complete linkage disequilibrium.

| Locus#1 | Locus#2 | P-Value | S.E. |
|-------------|-------------|-------------|-------------|
| mp1 | mp11 | 0.28 | 0.04 |
| mp1 | mp14 | 0.08 | 0.03 |
| mp1 | mp16 | 0.19 | 0.04 |
| mp1 | mp17 | 1.00 | 0.00 |
| mp1 | mp19 | 0.17 | 0.04 |
| mp1 | mp51 | 0.05 | 0.02 |
| mp1 | mp56 | 0.03 | 0.02 |
| mp1 | mp66 | 0.02 | 0.01 |
| mp1 | mp7 | 0.00 | 0.00 |
| mp1 | mp74 | 0.01 | 0.01 |
| mp1 | mp8 | 0.01 | 0.01 |
| mp11 | mp14 | 0.25 | 0.04 |
| mp11 | mp16 | 0.13 | 0.03 |
| mp11 | mp17 | 0.01 | 0.01 |
| mp11 | mp19 | 0.01 | 0.01 |
| mp11 | mp51 | 0.09 | 0.03 |
| mp11 | mp56 | 0.00 | 0.00 |
| mp11 | mp66 | 0.08 | 0.03 |
| mp11 | mp7 | 0.07 | 0.02 |
| mp11 | mp74 | 0.08 | 0.03 |
| mp11 | mp8 | 1.00 | 0.00 |
| mp14 | mp16 | 0.24 | 0.04 |
| mp14 | mp17 | 0.26 | 0.04 |
| mp14 | mp19 | 0.02 | 0.01 |
| mp14 | mp51 | 0.39 | 0.05 |
| mp14 | mp56 | 0.22 | 0.04 |
| mp14 | mp66 | 0.22 | 0.04 |
| mp14 | mp7 | 0.00 | 0.00 |
| mp14 | mp74 | 0.12 | 0.03 |
| mp14 | mp8 | 0.07 | 0.03 |
| mp16 | mp17 | 0.03 | 0.02 |
| mp16 | mp19 | 0.01 | 0.01 |
| mp16 | mp51 | 0.24 | 0.04 |
| mp16 | mp56 | 0.04 | 0.02 |
| mp16 | mp66 | 0.23 | 0.04 |
| mp16 | mp7 | 0.02 | 0.01 |
| mp16 | mp74 | 0.03 | 0.02 |
| mp16 | mp8 | 0.19 | 0.04 |
| mp17 | mp19 | 0.00 | 0.00 |
| mp17 | mp51 | 0.20 | 0.04 |

| | | | |
|-------------|-------------|-------------|-------------|
| mp17 | mp56 | 0.00 | 0.00 |
| mp17 | mp66 | 0.00 | 0.00 |
| mp17 | mp7 | 0.02 | 0.01 |
| mp17 | mp74 | 0.01 | 0.01 |
| mp17 | mp8 | 0.25 | 0.04 |
| mp19 | mp51 | 0.05 | 0.02 |
| mp19 | mp56 | 0.06 | 0.02 |
| mp19 | mp66 | 0.01 | 0.01 |
| mp19 | mp7 | 0.02 | 0.01 |
| mp19 | mp74 | 0.00 | 0.00 |
| mp19 | mp8 | 0.02 | 0.01 |
| mp51 | mp56 | 0.01 | 0.01 |
| mp51 | mp66 | 0.00 | 0.00 |
| mp51 | mp7 | 0.00 | 0.00 |
| mp51 | mp74 | 0.08 | 0.03 |
| mp51 | mp8 | 0.04 | 0.02 |
| mp56 | mp66 | 0.01 | 0.01 |
| mp56 | mp7 | 0.00 | 0.00 |
| mp56 | mp74 | 0.00 | 0.00 |
| mp56 | mp8 | 0.10 | 0.03 |
| mp66 | mp7 | 0.01 | 0.01 |
| mp66 | mp74 | 0.04 | 0.02 |
| mp66 | mp8 | 0.28 | 0.04 |
| mp7 | mp74 | 0.03 | 0.02 |
| mp7 | mp8 | 0.00 | 0.00 |
| mp74 | mp8 | 0.00 | 0.00 |

Table I-S4 The distribution of MLGs collected from each site over the nine DAPC groups.

| | G1 | G2 | G3 | G4 | G5 | G6 | G7 | G8 | G9 | Total |
|-------------|----|----|----|----|----|----|----|----|----|-------|
| Algerie_DJB | 0 | 4 | 2 | 0 | 0 | 2 | 1 | 0 | 7 | 16 |
| Algerie_K | 0 | 15 | 0 | 0 | 0 | 0 | 3 | 0 | 2 | 20 |
| Algerie_SK | 1 | 7 | 5 | 0 | 0 | 0 | 8 | 0 | 8 | 29 |
| Maroc_1 | 1 | 16 | 0 | 0 | 0 | 1 | 0 | 0 | 11 | 29 |
| Maroc_2 | 1 | 19 | 0 | 0 | 0 | 0 | 0 | 0 | 10 | 30 |
| Maroc_3 | 2 | 16 | 0 | 0 | 0 | 4 | 0 | 0 | 8 | 30 |
| Maroc_4 | 0 | 11 | 0 | 8 | 0 | 1 | 0 | 0 | 3 | 23 |
| Maroc_5 | 1 | 22 | 0 | 0 | 0 | 2 | 1 | 0 | 1 | 27 |
| Maroc_6 | 0 | 20 | 0 | 1 | 0 | 1 | 0 | 0 | 5 | 27 |
| Maroc_7 | 2 | 19 | 0 | 0 | 0 | 1 | 1 | 0 | 4 | 27 |
| Maroc_8 | 1 | 3 | 1 | 5 | 0 | 0 | 1 | 0 | 8 | 19 |
| Maroc_9 | 0 | 26 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 28 |
| Espagne_Maj | 0 | 21 | 0 | 0 | 0 | 2 | 2 | 0 | 7 | 32 |
| France_1 | 0 | 1 | 0 | 1 | 3 | 2 | 13 | 0 | 3 | 23 |
| France_2 | 1 | 1 | 3 | 1 | 1 | 0 | 17 | 0 | 4 | 28 |

| | | | | | | | | | | |
|--------------|----|---|----|---|----|----|----|----|---|----|
| France_3 | 0 | 2 | 0 | 0 | 0 | 1 | 14 | 0 | 4 | 21 |
| France_4 | 0 | 1 | 0 | 0 | 0 | 0 | 15 | 0 | 3 | 19 |
| France_5 | 1 | 0 | 0 | 0 | 1 | 0 | 17 | 0 | 4 | 23 |
| France_6 | 0 | 2 | 0 | 0 | 1 | 1 | 13 | 0 | 3 | 20 |
| FranceTrans | 0 | 6 | 0 | 0 | 0 | 2 | 34 | 0 | 2 | 44 |
| Italie_BR | 2 | 2 | 0 | 0 | 0 | 5 | 6 | 0 | 1 | 16 |
| Grece | 6 | 0 | 0 | 0 | 1 | 10 | 4 | 0 | 0 | 21 |
| Iran_1 | 1 | 0 | 0 | 0 | 28 | 3 | 5 | 0 | 1 | 38 |
| Iran_2 | 1 | 0 | 0 | 0 | 2 | 1 | 0 | 0 | 0 | 4 |
| Iran_4 | 1 | 0 | 0 | 0 | 2 | 0 | 1 | 0 | 0 | 4 |
| Iran_5 | 0 | 0 | 0 | 0 | 4 | 0 | 1 | 0 | 0 | 5 |
| Iran_6 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| Iran_7 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| Iran_8 | 1 | 0 | 0 | 0 | 2 | 0 | 3 | 0 | 0 | 6 |
| Ouzbekistan | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 1 | 5 |
| Turkmenistan | 1 | 0 | 0 | 0 | 3 | 2 | 0 | 0 | 1 | 7 |
| Turquie_01 | 5 | 1 | 0 | 0 | 0 | 15 | 1 | 0 | 3 | 25 |
| Turquie_05 | 12 | 0 | 0 | 0 | 7 | 8 | 0 | 0 | 0 | 27 |
| Turquie_08 | 14 | 1 | 0 | 0 | 3 | 3 | 0 | 0 | 0 | 21 |
| Turquie_14 | 8 | 0 | 0 | 0 | 2 | 8 | 6 | 0 | 1 | 25 |
| Turquie_16 | 1 | 0 | 0 | 0 | 28 | 0 | 0 | 0 | 0 | 29 |
| Turquie_19 | 3 | 0 | 0 | 0 | 12 | 9 | 2 | 0 | 3 | 29 |
| Turquie_20 | 3 | 4 | 0 | 0 | 7 | 7 | 0 | 0 | 7 | 28 |
| Syrie_1 | 9 | 1 | 0 | 0 | 2 | 4 | 0 | 0 | 0 | 16 |
| Syrie_10 | 10 | 0 | 0 | 0 | 1 | 6 | 0 | 0 | 3 | 20 |
| Syrie_2 | 0 | 1 | 2 | 0 | 0 | 2 | 2 | 6 | 2 | 15 |
| Syrie_3 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| Syrie_4 | 8 | 0 | 0 | 0 | 0 | 1 | 10 | 0 | 2 | 21 |
| Syrie_5 | 1 | 0 | 0 | 0 | 0 | 5 | 1 | 0 | 0 | 7 |
| Syrie_6 | 1 | 0 | 0 | 0 | 0 | 4 | 0 | 1 | 0 | 6 |
| Syrie_7 | 0 | 0 | 2 | 0 | 0 | 1 | 0 | 0 | 1 | 4 |
| Syrie_9 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 4 |
| HB_Bi | 0 | 1 | 2 | 0 | 0 | 3 | 0 | 2 | 0 | 8 |
| HB_Fe | 1 | 0 | 14 | 0 | 0 | 0 | 0 | 0 | 0 | 15 |
| HB_NA | 6 | 0 | 9 | 0 | 0 | 0 | 0 | 3 | 0 | 18 |
| HB_Ze | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 1 | 0 | 7 |
| HM_Am | 0 | 0 | 24 | 0 | 0 | 0 | 0 | 0 | 0 | 24 |
| HM_Bo | 2 | 1 | 0 | 0 | 0 | 3 | 5 | 0 | 1 | 12 |
| HM_Je | 5 | 0 | 10 | 0 | 0 | 13 | 0 | 0 | 1 | 29 |
| HM_Ma | 0 | 0 | 7 | 0 | 0 | 0 | 0 | 12 | 0 | 19 |
| HM_Rj | 0 | 0 | 8 | 0 | 0 | 0 | 2 | 0 | 0 | 10 |
| HN_Bc | 1 | 0 | 7 | 0 | 0 | 2 | 0 | 13 | 0 | 23 |
| HN_Bt | 0 | 0 | 20 | 0 | 0 | 0 | 0 | 0 | 0 | 20 |
| HS_DQ | 1 | 0 | 11 | 0 | 0 | 0 | 0 | 2 | 0 | 14 |

| | | | | | | | | | | |
|-----------|-----|-----|-----|----|-----|-----|-----|----|-----|------|
| HS_Ta | 0 | 0 | 5 | 0 | 0 | 2 | 0 | 4 | 0 | 11 |
| Awja | 2 | 0 | 2 | 1 | 0 | 19 | 0 | 5 | 3 | 32 |
| BM_Ha | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 6 |
| FB_D | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 3 |
| FB_G | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 0 | 7 |
| FB_L | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 5 |
| FN_WR | 0 | 0 | 7 | 0 | 0 | 0 | 0 | 2 | 0 | 9 |
| NN_Ij | 0 | 0 | 0 | 13 | 0 | 0 | 0 | 0 | 0 | 13 |
| NN_Kf | 0 | 0 | 0 | 10 | 0 | 0 | 0 | 0 | 0 | 10 |
| IN_Ba | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 7 | 0 | 8 |
| IN_Be | 3 | 0 | 9 | 0 | 0 | 2 | 0 | 2 | 3 | 19 |
| IN_WR | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 14 | 0 | 15 |
| Istanbuli | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| KB_Ba | 12 | 2 | 0 | 0 | 0 | 1 | 2 | 0 | 9 | 26 |
| KB_Bi | 14 | 2 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 20 |
| KB_EM | 17 | 0 | 1 | 0 | 0 | 3 | 2 | 0 | 7 | 30 |
| KB_Ir1 | 4 | 2 | 0 | 0 | 0 | 5 | 5 | 0 | 0 | 16 |
| KB_Ir2 | 2 | 3 | 0 | 0 | 1 | 5 | 2 | 0 | 2 | 15 |
| KB_KL | 22 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 5 | 28 |
| KB_Ko | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 5 |
| KB_La | 14 | 2 | 0 | 0 | 0 | 3 | 1 | 0 | 5 | 25 |
| KB_Lu | 19 | 1 | 0 | 0 | 1 | 2 | 3 | 0 | 2 | 28 |
| KB_Sh | 11 | 0 | 2 | 0 | 0 | 3 | 0 | 0 | 1 | 17 |
| KM_Ma | 1 | 4 | 2 | 0 | 0 | 0 | 4 | 0 | 0 | 11 |
| KN_As | 0 | 3 | 1 | 0 | 0 | 7 | 1 | 0 | 3 | 15 |
| KN_EQ | 6 | 1 | 1 | 0 | 0 | 5 | 2 | 0 | 1 | 16 |
| KN_Mi | 0 | 0 | 0 | 0 | 0 | 7 | 1 | 3 | 1 | 12 |
| KS_Ba | 3 | 2 | 0 | 0 | 0 | 18 | 1 | 0 | 6 | 30 |
| KS_Bl | 4 | 0 | 1 | 0 | 0 | 22 | 1 | 0 | 1 | 29 |
| KS_Ch | 8 | 1 | 0 | 0 | 0 | 8 | 0 | 0 | 4 | 21 |
| KS_MJ | 1 | 0 | 2 | 0 | 0 | 10 | 2 | 0 | 2 | 17 |
| KS_Rm | 0 | 0 | 2 | 0 | 0 | 3 | 1 | 0 | 0 | 6 |
| KS_Ta | 2 | 1 | 0 | 0 | 0 | 11 | 0 | 2 | 1 | 17 |
| Total | 265 | 250 | 172 | 41 | 118 | 274 | 236 | 85 | 184 | 1625 |

Table I-S5 Allele's loading over the eight DAPC axes.

| | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 | PC7 | PC8 |
|-------------|------|------|------|------|------|------|------|------|
| mp16.125 | 0.09 | | | | | | | |
| mp51.159 | 0.07 | | | | | | | |
| mp51.160 | 0.29 | 0.73 | | | | | | |
| mp7.230 | 0.05 | | | | | | | 0.06 |
| mp1.100 | | | 0.07 | | | | 0.17 | |
| mp7.234 | | | 0.09 | 0.17 | | | | |
| mp16.105 | | | | 0.12 | | | | |
| mp17.183 | | | | 0.09 | | | | |
| mp11.154 | | | | | 0.06 | | | 0.08 |
| mp66.138 | | | | | 0.08 | | | |
| mp66.153 | | | | | 0.11 | | | |
| mp51.157 | | | | | | 0.10 | | 0.12 |
| mp7.238 | | | | | | 0.10 | | |
| mp19.179 | | | | | | | 0.07 | 0.05 |
| mp19.177 | | | | | | | 0.08 | |
| mp56.137 | | | | | | | 0.12 | |
| mp8.103 | | | | | | | 0.05 | |
| mp7.236 | | | | | | | | 0.11 |
| mp74.156 | | | | | | | | 0.09 |
| variability | 0.37 | 0.23 | 0.11 | 0.09 | 0.07 | 0.05 | 0.04 | 0.03 |

Additional file 2

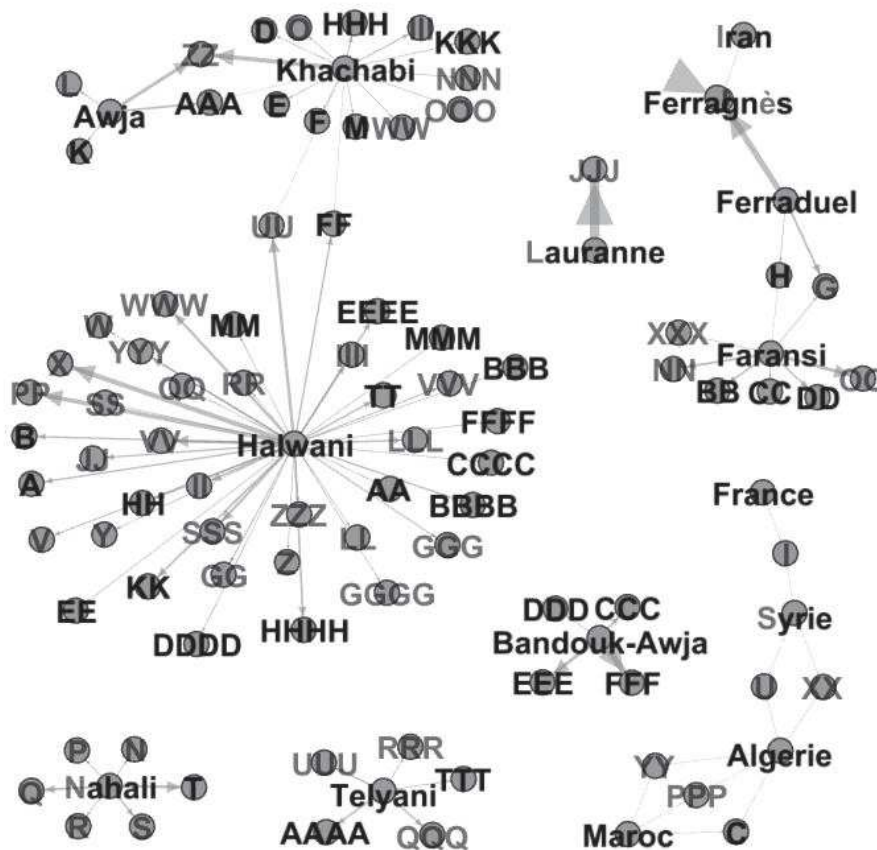


Figure I-S1 Network-like representation for clones identified within our samples. Each clone is represented by a tie linking an alphabet node corresponding to the MLGs to a node of each region and/or cultivar having common or specific redundant MLGs. The thickness of the tie corresponds to number of redundancy.

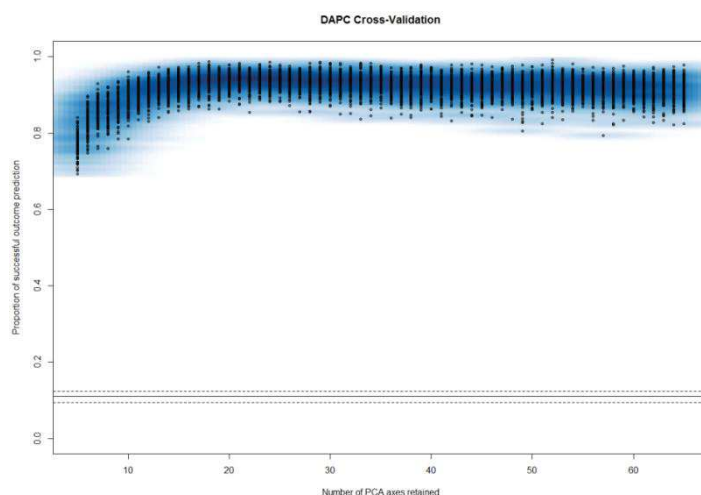


Figure I-S2 Cross-validation procedure to choose the optimal number of Principal Components for the DAPC analysis. the optimal number of Principal Components was 19 PCs.

« L'identité n'est pas donnée une fois pour toutes, elle se construit et se transforme tout au long de l'existence. »

Amin Maalouf, 1998
Les identités meurtrières

Chapitre 2

Taxonomie locale et caractérisation de la diversité de l'amandier cultivé au Liban

(En préparation pour la revue 'Revue d'Ethnoécologie')

Auteurs

Bariaa Hamadeh

Biologie des populations, Doctorante, Université Libanaise, Ecole Doctorale des Sciences et des Technologies, Hadath, Liban.

Université de Montpellier, Ecole Doctorale GAIA (N°584), Place Eugène Bataillon ; Bâtiment 13 / 3^{ème} étage, 34095 Montpellier Cedex 5.

Institut de Recherche Agronomique Libanais, Fanar, Liban.

CIRAD - UMR AGAP, F-34398 Montpellier, France. * Adresse actuelle.

AGAP, Univ. Montpellier, CIRAD, INRA, Montpellier SupAgro, Montpellier, France

bariaa.hamadeh@gmail.com

Hélène I. Joly

Généticienne, Chargé de Recherche, CIRAD - UMR AGAP, F-34398 Montpellier, France.

AGAP, Univ. Montpellier, CIRAD, INRA, Montpellier SupAgro, Montpellier, France.

helene.joly@cirad.fr

Pascale Moity-Maizi

Anthropologue, Maître de Conférence, SupAgro - UMR GRED, Institut des Régions Chaudes (IRC) - 5098 Montpellier cedex 05, France.

pascale.maizi@supagro.fr

Résumé

L'interaction entre les différents membres d'une société et les plantes qu'ils cultivent est mise en évidence par la diversité des connaissances et des termes utilisés pour décrire la diversité variétale de l'espèce concernée et l'ancienneté de sa culture. Nous avons identifié la complexité et les critères de classification de l'amandier cultivé au Liban à partir d'entretiens réalisés auprès de 65 informateurs rencontrés dans tout le pays. Les résultats de l'analyse de listes libres a permis d'identifier 24 types nommés d'amandiers. Cette étude souligne une dichotomie essentielle dans la nomenclature locale sur deux caractéristiques : 1) amer (Murr) / doux (Helou) (la littérature indique que ce trait pourrait être le 1^{er} caractère de domestication), 2) coque dure (Khachabi) / coque tendre (Ferk). L'analyse de l'abondance des variétés dans les vergers et de l'occurrence des types nommés dans les listes libres permet aussi de montrer que 'Khachabi' et 'Halwani' sont deux types nommés que l'on qualifie de « phares », vu leur importance sur un double plan culturel et agronomique. La taxonomie locale de l'amandier au Liban est profonde, elle met en évidence une diversité importante sur le plan variétal et une diversité intra-variétale pour certains types nommés.

Mots clés : amande, type nommé, cultivar, taxonomie, classification locale, Liban.

Introduction

La dynamique de la diversité des espèces cultivées est influencée par des processus biologiques et socio-culturels. Cependant, pour les biologistes l'évaluation de cette diversité s'est longtemps reposée sur des descripteurs morpho-physiologiques ou agronomiques (en rapport avec la sélection et l'utilisation des plantes cultivées) et génétiques (en utilisant différents types de marqueurs enzymatiques, AFLP, microsatellites, etc...). Aujourd'hui, on admet qu'il est nécessaire de prendre aussi en compte des dimensions sociales par une approche anthropologique permettant de recueillir les classifications locales et de construire une taxonomie locale (Berlin 1992) ; ceci permet de comprendre la diversité des plantes cultivées telle qu'elle est perçue, conçue et maîtrisée par les agriculteurs eux-mêmes.

On appelle taxonomie locale un système de nomenclature et de classification développé par une société, lui permettant de décrire et de classer mais aussi de transmettre des éléments significatifs de son environnement et notamment les plantes. Tout processus de classification repose sur trois mécanismes (Friedberg 1991): 1) l'observation qui conduit à l'identification des caractéristiques de la plante, 2) la dénomination qui permet d'associer des lexèmes aux différentes caractéristiques identifiées, 3) la classification qui permet de regrouper la plante identifiée avec d'autres plantes qui partagent, selon les circonstances¹, les mêmes caractéristiques du point de vue de ceux qui les distinguent ou les décrivent. La nomenclature repose sur un lien étroit entre l'objet et le nom donné (lexème) qui donne une signification à cet objet. La catégorisation définit et exprime les relations entre les êtres vivants, aide à l'apprentissage et à la communication, sert de dispositif cognitif et sémantique pour stocker et extraire l'information (Berlin 1992). Berlin (1992) a proposé des principes de

¹ C'est ce qui explique l'évolution continue des classifications botaniques par exemple, au fur et à mesure que les circonstances et connaissances scientifiques se transforment. Voir Bonneuil and Thomas, 2009.

catégorisation utilisés pour décrire l'arrangement et le contenu des classifications vernaculaires, qui prennent souvent la forme de taxonomies.

Les taxonomies de plantes sont dites structurées en profondeur quand plusieurs lexèmes peuvent être reliés pour décrire le niveau intra-spécifique. Enfin, il est aujourd'hui reconnu qu'un système de classification et le niveau de complexité de la taxonomie locale servent aussi d'indicateurs pour évaluer la diversité intra-spécifique des espèces cultivées (Brush 2004). Dans chaque société, le nom donné à une plante cultivée résulte d'observations, mais aussi d'usages, d'échanges, de transferts de connaissances ou de matériaux, qui se succèdent dans l'histoire des agriculteurs.

La taxonomie locale intra-spécifique a été étudiée notamment chez le Manioc (*Manihot esculenta*) en Guyane (Elias *et al.* 2001), le maïs (*Zea mays*) au Mexique *et al.* 2005, Brush & Perales 2007), le taro (*Colocasia esculenta*) au Vanuatu (Caillon *et al.* 2006) et le Sorgho (*Sorghum bicolor*) en Ethiopie et au Cameroun (Mekbib 2007, Barnaud *et al.* 2007). Plus récemment, les taxonomies locales pour certaines espèces fruitières méditerranéennes ont été reconstruites : pour le figuier au Maroc (Hmimsa *et al.* 2012), l'olivier au Maroc (Haouane 2012). Ces études ont permis d'identifier les connaissances de différents groupes sociaux à propos de la diversité intra-spécifique de ces espèces fruitières cultivées. Elles ont également montré que la nomenclature vernaculaire est en lien direct avec l'usage de la plante mais qu'elle se construit aussi sur des critères morphologiques tels que la couleur, la pilosité, etc. Avec ces auteurs, nous considérons que l'importance de la diversité intra-spécifique peut être expliquée par une caractéristique morphologique particulière, par des usages spécifiques, par une origine particulière, par l'existence de mythes, etc. Barnaud *et al.* (2009) ont démontré par exemple, en étudiant le sorgho au Cameroun dans la société Duppa, que certains types nommés renvoyaient directement à des usages rituels. Ces études soulignent que les termes désignant les plantes cultivées font souvent référence à des lieux d'origine ou à des personnes

qui sont à l'origine de l'introduction de la variété, ce que l'on ne trouve pas en général dans les taxonomies concernant les espèces spontanées. Hmimsa *et al.* (2012) ont démontré que les usages ne sont pas exprimés dans le nom des variétés de figuier cultivé au Maroc, bien que les agriculteurs les différencient et les cultivent en fonction de divers usages.

Le Proche-Orient est considéré comme le centre de diversification des arbres fruitiers méditerranéens (Zohary *et al.* 2012). L'amandier (*Prunus dulcis* Mill.), est un arbre fruitier méditerranéen cultivé au Liban depuis des siècles. La diversité intra-spécifique des arbres fruitiers a été peu étudiée dans cette région, on dispose cependant de l'inventaire réalisé par les services du général Gouraud en 1921 effectué en Syrie et en Cilicie (Services de général Gouraud 1921). Il mentionne l'amandier cultivé et la présence de trois variétés. Depuis, les sociétés du Moyen-Orient ont connu de nombreux bouleversements économiques et politiques, marqués par des échanges et flux réguliers de populations entre pays, du fait de guerres mais aussi du fait d'échanges économiques plus nombreux. La circulation des groupes humains, de matériel végétal et de connaissances s'est accélérée. Actuellement, nous avons observé que l'amandier est cultivé dans des agroécosystèmes traditionnels et dans des vergers commerciaux. Chalak *et al.* (2007) ont ainsi identifié 18 variétés d'amandier et ont caractérisé la diversité intra-spécifique de l'amandier cultivé au Liban en se basant sur des descripteurs morphologiques et moléculaires. Delplancke *et al.* (2012) a aussi montré récemment l'importante diversité génétique de l'amandier cultivé au Proche-Orient.

Dans notre étude, nous avons cherché à décrire les processus de catégorisation appliqués à l'amandier au Liban afin de reconstruire la taxonomie locale et de comprendre la place de l'amandier dans ce pays, en allant au-delà de son expansion spatiale ou de ses usages alimentaires. Nous avons donc fait un inventaire des termes utilisés – reflets de connaissances locales - concernant l'amandier cultivé au Liban et sa variabilité afin de dévoiler le système

de catégorisation socio-culturelle et cognitive locale à différentes échelles du Liban (pays, régions et village).

Notre démarche mobilise des outils de l'anthropologie, associés à un suivi des vergers d'amandiers, à la construction de listes libres et à une approche sémantique. L'évaluation de la diversité cognitive portant sur l'amandier au Liban s'est essentiellement appuyée sur l'analyse des listes libres à l'échelle du pays avec un focus à l'échelle du village emblématique de Zeghrine, choisi car la culture de l'amandier y est ancienne et toujours très dynamique.

Les enquêtes et observations de terrain à l'échelle du pays ont été réalisées en 2010 et 2011 et les investigations plus ciblées dans le village de Zeghrine se sont déroulées en 2014. Ces missions de terrain ont toutes été réalisées dans le cadre des recherches conduites par l'Institut de Recherche Agronomique Libanais, sous la tutelle du Ministère de l'Agriculture.

Sites d'étude et protocole d'échantillonnage

Dans un premier temps, nous avons visité les sites déjà explorés dans les études précédentes sur l'amandier au Liban (Post 1932, Mouterde 1966, Chalak *et al.* 2007). Pour identifier de nouveaux vergers, une liste de contact des municipalités a été téléchargée à partir du site internet du Ministère d'Intérieur et des Municipalités. Nous avons établi un premier contact avec toutes les mairies du pays pour nous informer de l'existence de plantations d'amandiers de plus de 25 ans dans chaque village (l'âge du verger est précisé afin de faciliter l'identification d'informateurs expérimentés dans la culture de l'amandier). Nous avons aussi suivi un itinéraire défini au préalable sur carte ou photographie aérienne pour identifier les villages sans mairie où il pouvait aussi y avoir des vergers ; nous nous sommes arrêtés dans chaque verger d'amandier repéré. Au total nous avons retenu 87 vergers dans 73 villages.

Pour chaque verger nous avons cherché à identifier un informateur correspondant à la personne responsable ou considérée comme la plus compétente sur chaque exploitation. Ces

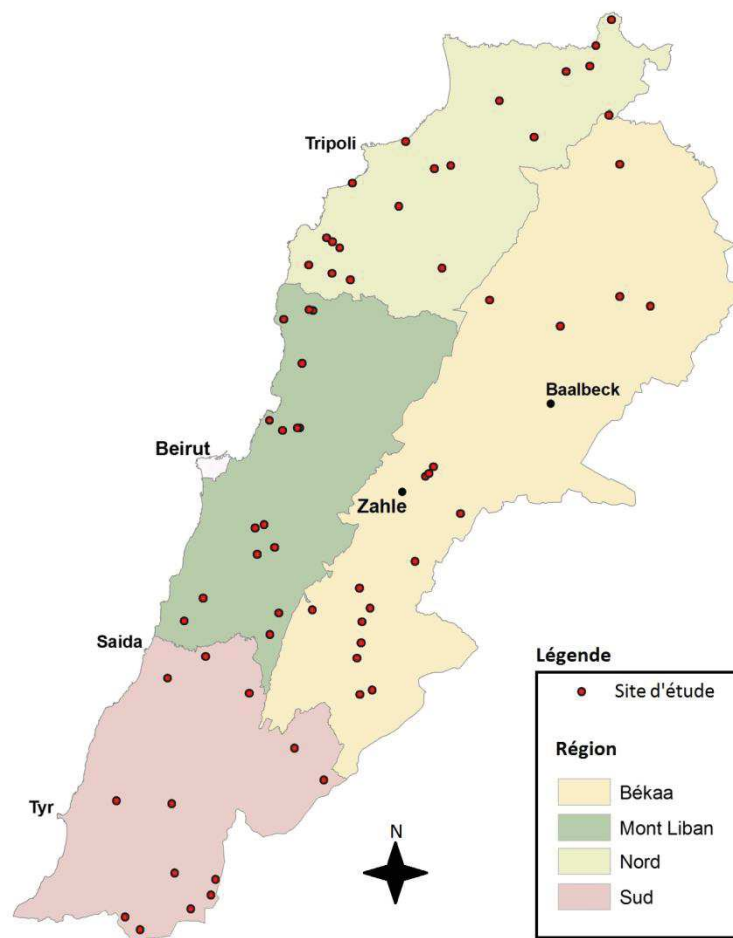


Figure II-1. Carte du Liban représentant la distribution géographique des 65 sites de référence.

personnes ont parfois pu être identifiées au hasard de rencontres préalables lors des collectes de matériel végétal. Mais pour 22 vergers, aucun informateur n'a pu être identifié. Ces vergers ont donc été retirés du jeu de données qui comprend donc au final 65 vergers dotés d'informateurs et localisés dans 54 villages (Figure II-1).

Ce sont nos premières prospections et nos échanges préalables avec des agriculteurs qui nous ont permis de noter la particularité du village de Zeghrine à l'échelle du pays. Les agriculteurs de Zeghrine sont en effet des connaisseurs de la diversité variétale et cultivent l'amandier depuis longtemps. Cette culture est aussi la base de leur revenu. Ce village, rattaché à la commune de Baalbek-el-Hermel, dans la Bekaa Nord, a donc été retenu pour une étude approfondie à l'échelle locale. Treize informateurs ont été retenus pour réaliser l'étude à Zeghrine.

Déroulement des enquêtes

Les 65 informateurs identifiés pour l'étude à l'échelle du pays et les 13 informateurs identifiés à l'échelle du village de Zeghrine ont fait l'objet d'une enquête individuelle par questionnaire (Annexe 1). Ces informateurs-experts sont soit des propriétaires agriculteurs, soit des propriétaires non agriculteurs, ou encore des techniciens agricoles ou des personnes de référence pour l'histoire des vergers de la collectivité locale. Nous n'avons donc pas interrogé dans cette étude les pépiniéristes, les membres d'ONG ni les représentants du ministère de l'agriculture qui peuvent aussi avoir une expertise sur les amandiers. Nous avons privilégié ici des acteurs agricoles proches du terrain, dotés d'une forte reconnaissance auprès de la population locale.

Nos questions concernaient les pratiques culturelles, les connaissances de la diversité variétale, les critères de choix et l'origine du matériel végétal, l'état de la plantation, le type de verger (verger commercial ou verger traditionnel familial), enfin les variétés cultivées. Une

attention particulière a été accordée à la dénomination de ces dernières et à leur caractérisation par les informateurs.

Nous avons aussi réalisé une liste libre (Martin 2004) en interrogeant les mêmes informateurs (N=65 et N=13), sur la question suivante : « quelles variétés d'amandier cultivées connaissez-vous au Liban, même si vous ne les avez pas dans vos vergers ? ». Chaque individu énonce alors une liste de noms enregistrés dans l'ordre de leur énonciation, ce qui permet de révéler les types nommés les plus connus de chaque informateur. Ensuite, chacun d'eux est interrogé sur les types qu'il a nommés, l'objectif étant alors de relever des indicateurs de l'ancrage historique et social ainsi que des caractéristiques biologiques de chaque type nommé. Nos informateurs ont pu aussi préciser les critères de différenciation entre ces types nommés. Les critères morphologiques et agronomiques énoncés pour caractériser les types nommés ont été systématiquement relevés auprès de chaque informateur. Ils nous ont ainsi permis d'identifier ensuite des synonymes et homonymes entre types et de limiter les problèmes d'analyse liés aux répétitions.

Le traitement des questionnaires et des entretiens

Tous les discours ont été recueillis en arabe puis traduits en français. Ils ont été enregistrés dans une base de données anthropologiques, indexés selon les pratiques et selon des thèmes se rapportant aux relations sociales, aux valeurs culturelles ou économiques liant les libanais et les amandiers de ce pays.

Une approche sémantique a été réalisée pour étudier la signification des mots et les rapports de sens entre les mots (homonymie) : ce sont les désignations ou signifiés, qui nous ont permis d'élaborer une nomenclature variétale de l'amandier à l'échelle du Liban. Selon Berlin (1992), l'analyse linguistique est essentielle pour les études taxonomiques qui révèlent les systèmes de classification traditionnels. D'une part, le langage est la « porte d'accès aux

connaissances et aux perceptions» et d'autre part, beaucoup d'informations sont transmises à l'aide de représentations symboliques (Cotton 1996).

Nous avons aussi construit une analyse quantitative permettant d'évaluer à la fois la diversité de l'amandier cultivé et son importance culturelle: la liste libre, basée sur la mémoire, repose en effet sur l'hypothèse que chaque informateur a tendance à citer les termes qu'il connaît en suivant un ordre correspondant à son degré de familiarité avec l'objet désigné. Les termes les plus souvent mentionnés représentent alors les variétés les plus connues. Les types nommés les plus cultivés et les plus cités sont désignés ici par le terme « types nommés phares ». Une comparaison a donc aussi été effectuée entre la fréquence des types nommés et leur abondance effective dans les vergers étudiés.

Le calcul de la diversité cognitive a été réalisé grâce à l'indice de Sutrop (Sutrop 2001) implémenté dans le logiciel Flame v1.2 (Pennec *et al.* 2012). Cet indice permet d'établir un score d'importance culturelle en combinant deux critères: la fréquence de citation et l'ordre d'apparition de l'item dans la liste. L'indice de Sutrop a été utilisé dans cette étude pour avoir une première approche de la valeur culturelle des variétés d'amandier cultivées dans la société libanaise.

Description des informateurs

Sur les 65 informateurs rencontrés à l'échelle du pays, six sont des techniciens agricoles non propriétaires de vergers, tous les autres sont des agriculteurs propriétaires de leurs vergers. Parmi les agriculteurs, deux sont des femmes de la région du Liban nord. Sur les 13 informateurs du village de Zeghrine, deux sont des techniciens agricoles non propriétaires de vergers, les autres sont des agriculteurs propriétaires de leurs vergers.

1 **Tableau II-1.** Nombre de listes libres collectées par âge du verger.

| | > 40 | [40 - 20[| [20 - 10[| Variable |
|------------------------------|------|-----------|-----------|----------|
| Nombre de listes | 11 | 18 | 21 | 6 |
| Longueur moyenne des listes | 3 | 3 | 3 | 4 |
| Nombre de types nommés | 10 | 14 | 16 | 12 |
| Nombre total de types nommés | 36 | 51 | 60 | 27 |

2

3

4 **Tableau II-2.** Nombre de vergers échantillonnés dans les quatre zones éco-

5 **géographiques.**

| Etat du verger | Type du verger | Mont Liban | Liban Nord | Bekaa | Liban Sud | Total |
|------------------|----------------|------------|------------|-------|-----------|-------|
| Entretenu | Familial | 9 | 1 | 3 | 2 | 15 |
| | Commercial | 6 | 10 | 9 | 5 | 30 |
| Négligé | Familial | 2 | 0 | 2 | 0 | 4 |
| | Commercial | 0 | 1 | 0 | 0 | 2 |
| Abandonné | Familial | 0 | 6 | 2 | 4 | 12 |
| | Commercial | 1 | 0 | 0 | 1 | 2 |

Description des vergers

Nous avons interrogés 21 informateurs dont les vergers sont récents (10 à 20 ans), 18 dont les vergers ont de 20 à 40 ans, 18 dont les vergers ont plus de 40 ans, et 6 dont les vergers ont des plants d'âges variables (Tableau II-1). Ces derniers sont souvent d'anciens vergers rénovés qui constituent une source de revenus importante pour les informateurs.

L'amandier est cultivé dans 32 vergers familiaux, dans 31 vergers commerciaux dont la production est destinée au marché Libanais, et dans deux anciens vergers commerciaux qui sont devenus familiaux (Tableau II-2). Sur les 65 informateurs, 45 entretiennent leurs vergers, 14 les ont abandonné (12 vergers familiaux et 2 anciens vergers commerciaux), et six informateurs consomment ou vendent leur production d'amandes mais n'effectuent aucun entretien (vergers négligés) (Tableau II-2).

Dénomination et critères de classification de l'amandier au Liban

La désignation des variétés est généralement exprimée par un terme de base qui peut-être un monolexème (par exemple : 'Khachabi', 'Halwani', 'Awja', etc...) ou un dilexème (par exemple : 'Oum Omar', 'Cheikh-el-achiri', 'Bandouk Awja') (Tableau II-3). Parfois un déterminant s'ajoute pour qualifier certaines variétés relevant de la même désignation. Les informateurs utilisent un ou deux critères pour caractériser un type nommé puis ils lui attribuent un nom (Tableau II-3). Tous les termes de base donnés sont justifiés en se référant à des critères de classification. Les termes utilisés pour caractériser ou hiérarchiser les types nommés renvoient à trois grandes classes de critères : i/ morphologique et organoleptique, ii/ origine, iii/ métaphore (Tableau II-3 et 4).

Tableau II-3. Critères de classification du terme de base.

| Classe | Critère de classification | Occurrence (%)* | Noms cités |
|--|------------------------------------|------------------------|---|
| Morphologique ou organoleptique | Dureté de la coque | 35 | Ferk, Khachabi, Chakouchi, Kassi, Klaibi, Maksar, Nos-Ferk, Nos-Khachabi |
| | Forme et taille du fruit | 15 | Awja, Mdaabal |
| | Gout de l'amandon | 13 | Morr, Halwani |
| | Pilosité du fruit | 4 | Abrach, Mallissi, Bou Waber |
| Origine | Pays ou région | 18 | Baladi, Telyani/Itali, Istanbouli, Faransi, Afsaji, Ajami, Masri, Amercai |
| | Nom de la personne qui l'a propagé | 11 | Nahali, Oum Omar, Oum Rachid, Oum Hsein |
| Métaphore | Métaphores | 5 | Cheikh-el-achiri, Bandouk Awja, Barri, Jawwi, Laouz |

*Pourcentage d'occurrence de noms vernaculaires associés à chaque critère de classification.

Tableau II-4. Critères de classification du déterminant.

| Terme de base | Classe du déterminant | Critère de classification du déterminant | déterminant |
|----------------------|---------------------------------|---|--|
| Halwani | Morphologique ou organoleptique | Nombre d'amandon | Lebb, Lebbein |
| Faransi, Ferk | | Forme et taille du fruit | Abrach, Mallissi, Tawil, Mabroum/Mdaabal |
| Ferk, Khachabi | | Gout de l'amandon | Helou, Morr |
| Ferk | Origine | Pays | Souri |

La première classe (67% d'occurrence) a pu être organisée en distinguant 4 sous-groupes de critères :

- 1) Dureté de la coque (35% d'occurrence) : ce critère a été désigné par les termes 'Khachabi' (Khachab : bois, à coque dure comme du bois) et 'Ferk' (friable, à coque tendre). Les informateurs identifient un gradient de dureté de la coque qui passe de la coque dure à la coque tendre avec la présence de types nommés dont la coque a une dureté intermédiaire: 'Nos-Khachabi' (demi-Khachabi) et 'Nos-Ferk' (demi-Ferk).
- 2) La forme et la taille du fruit (15% d'occurrence) : ce critère est relatif à la morphologie du fruit entier (l'amande) et du noyau (l'amandon). Le nom 'Awja' (courbée ou incurvée) indique la forme du fruit. Les termes 'Mabroum' ou 'Mdaabal' traduisibles par « rond » indiquent aussi la forme du fruit.
- 3) Critère organoleptique (13% d'occurrence): Le nom 'morr' (Amer), donné pour indiquer l'amertume des amandons, fruits des arbres sauvages² ou féroces (semis de plants domestiqués, cultivés, retournés à l'état sauvage). Le terme 'Halwani' qui veut dire « sucré » fait référence à la douceur de l'amandon ; il est aussi utilisé pour désigner le pâtissier, ce qui renvoie indirectement à l'usage de cette variété dans les pâtisseries.
- 4) La pilosité du fruit (4% d'occurrence): le terme 'Mallisi' souligne la surface lisse du fruit. Par opposition, 'Abrach' ou 'Bou Waber' sont deux termes traduisibles par « velu », indiquant la présence de poils sur le fruit. 'Aghbar' est un synonyme des termes précédents, traduisible par « poussiéreux », qui réfère aussi à l'écorce velue du fruit.

La classe relative aux origines (29 % d'occurrence) est organisée en deux groupes :

- 1) Le groupe des termes relatifs à des lieux (18% des occurrences) : il comprend les termes génériques qui différencient le « local » (tout ce qui est local – et national- s'appelle

² On ne peut savoir s'il s'agit d'arbres sauvages ou spontanés puisqu'on ne pourrait pas encore préciser l'origine de l'amandier domestiqué.

'Baladi') de ce qui provient d'ailleurs. Ainsi, par exemple : 'Faransi' pour la France, 'Itali/Telyani' pour l'Italie, 'Istanbuli' pour la Turquie, et 'Amercani' pour une provenance supposée de l'Amérique.

2) Le groupe des termes qui désignent une personne qui aurait propagé la variété (11 % d'occurrence) : 'Nahali' (le mot veut dire l'apiculteur, mais les informateurs relient ce terme au nom même de la personne qui l'a introduite, à savoir « le père Nahli de Kfarhatna » un village à Batroun, Liban Nord) ; 'Oum Omar', 'Oum Rachid' et 'Oum Hsein' (Oum : maman, Omar, Rachid et Hsein sont des prénoms masculins; et une femme est plutôt appelée « maman de ... » avec le prénom du premier descendant masculin. Ces types sont nommés avec le nom de la femme qui aurait introduit ou propagé le type nommé). Cette catégorie de termes reflète donc l'importance d'un réseau social ou même d'une lignée de diffusion et d'adoption de certaines variétés sur un lieu d'origine.

La classe des métaphores (4% d'occurrence) est organisée en deux groupes :

1) Les qualificatifs dévalorisant ou au contraire valorisant la variété sont incluses ici des :

- métaphores partielles de type : 'Bandouk Awja' (traduisible par : « batard de 'Awja' »), la forme du fruit est proche de celle de la variété 'Awja' mais les fruits sont plus petits. L'agriculteur de la région de Hallat qui nous a livré ce qualificatif a indiqué que cette variété résulte de l'hybridation locale entre la variété 'Awja' et une autre variété inconnue ; selon cet agriculteur, « la variété inconnue est probablement 'Halwani' car la présence de 'Bandouk Awja' avec cette dernière donne une bonne quantité de fruit ».
- métaphores complètes : par opposition à 'Bandouk-Awja', le qualificatif 'Cheikh-el-achiri', est particulièrement valorisant (Cheikh : chef, achiri : le clan), car il désigne le statut social le plus prestigieux, qui impose le respect. La variété qui est désignée ainsi est de très grand calibre (12 cm pour le fruit) et a beaucoup de valeur pour la société du Nord; elle est aussi particulièrement prisée pour la consommation en fruits verts. Selon l'agricultrice interrogée à

Wadi el Rihan au Liban nord « On pense aussi qu'elle est anciennement sélectionnée et cultivée dans cette région mais nous n'avons pas d'information sur le temps depuis lequel cette variété est présente ». Dans d'autres régions, seule une personne a nommé cette variété parce qu'elle en a entendu parler. Nous constatons aussi que ce type nommé est essentiellement cultivé dans le Nord.

2) Le second groupe de cette classe comprend des termes qui désignent des parties de l'être humain ou des éléments de la nature :

Le terme 'Laouz' (avec des variantes phonétiques : 'Luz', 'Laouzé', 'Laouzi') qui veut dire amygdale ou amygdaloïde (en Français « forme d'amande ») désigne aussi bien le fruit que l'arbre ; et le terme 'Mkhaddad' (qui signifie « avec des joues », et par extension tout être avec « une face rouge quand il mûrit ») est mentionné uniquement dans les vergers de Zeghrine.

La métaphore 'Jawwi' qui évoque l'air est affectée à tout ce qui est domestiqué, à toute variété qui n'existe pas à l'état sauvage. Par opposition, on utilise le terme 'Barri' (Barr: « terre », « barri » : en lien avec la terre) pour tout ce qui est présent à l'état sauvage, spontané ou rustique.

Les synonymes peuvent apporter une information supplémentaire pour qualifier un type. Par exemple, le type nommé 'Khachabi' compte six synonymes dont 4 termes relatifs à la dureté de la coque : 'Kassi' (dure), 'Klaibi' (en lien avec les chiens, faisant allusion à la puissance de la mâchoire pour casser la coque), 'Kasr' (« à casser »), 'Chakouchi' (Chakouch: Marteau) qui sous-entend le besoin d'un marteau pour casser la coque. Ce dernier terme est mentionné six fois dans le nord et une seule fois dans la Bekaa chez un agriculteur qui est en relation avec des agriculteurs dans le Nord. Enfin, le terme Barri (sauvage) indique la présence de ce type dans la nature, et 'Baladi' (« du pays ») désigne clairement le caractère local du type nommé 'Khachabi'. Par opposition, les termes 'Ferk' et/ou 'Halwani' désignent différentes variétés à coque tendre. Mais comme la variété 'Halwani' est la plus répandue dans les différentes zones

géographiques, les gens utilisent aussi bien le nom 'Ferk' que 'Halwani' pour parler des variétés cultivées à coque tendre, dominantes dans le pays.

D'autres termes sont employés pour ajouter un niveau de précision supplémentaire et mieux différencier ainsi des variétés cultivées. Ce sont des critères morphologiques et organoleptiques tels que la forme et la taille du fruit ('Mabroum'/ 'Mdaabal' : rond ; ou tawil : long- pour dire de grande taille - ; et zghir : petit), le nombre d'amandon (lebb : un seul amandon, lebbein : deux amandons), le gout de l'amandon (morr : amer ; helou: doux). Ce sont aussi des critères liés cette fois à l'usage ('Ferk akhdar': se consomme en fruit vert), et à l'origine ('Ferk souri': Ferk provient de la Syrie).

D'autres critères liés aux usages ont été identifiés lors des enquêtes par entretiens. Ce sont les suivants:

- 1) Usage médicinal : concerne essentiellement les amandes amères qui sont utilisées en coques dans des tisanes contre le rhume.
- 2) Usage agronomique : 'Khachabi', adaptée aux conditions phytopathologiques et pédoclimatiques du pays, est utilisée comme porte-greffe quand elle porte des amandes amères pour greffer les variétés à coque tendre moins résistantes aux ravageurs. Nous avons rencontré des agriculteurs utilisant même la variété 'Khachabi' à amande douce comme porte-greffe. Selon l'agriculteur de Zouk Mikael dans le Mont Liban : «la variété 'Khachabi' est une variété locale présente dans le pays depuis longtemps, même les arbres à amande douce sont plus résistants aux conditions climatiques et aux ravageurs, c'est pourquoi je les utilise comme porte-greffe».
- 3) Usage alimentaire : toutes les amandes produites au Liban sont de fait consommées dans le pays. Nos informateurs identifient 'Khachabi', comme ayant une meilleure aptitude à la conservation pendant l'hiver grâce à sa coque dure. La variété

'Halwani' à coque tendre est au contraire destinée à une consommation du fruit vert (le fruit est alors consommé entier lorsque l'amandon est encore gélifié), ou de l'amandon frais (à maturité de l'amandon) et sec (amandon sec).

4) Usage marchand : la variété 'Halwani lebbein' est particulièrement valorisée sur le marché du fait de sa précocité et de la présence d'amandons doubles. Selon un agriculteur de Tanbourite au Sud du Liban : «la variété 'Halwani' est la première variété présente sur le marché en vert, les 200g se vendent 7000 L.L.» (3 euro/200g)³. Un autre agriculteur à Sfarai dans la région du Liban Sud ne cultive que le 'Halwani lebbein' (à amandon double) car «ce caractère est recherché et très apprécié par les consommateurs libanais».

Analyse des listes libres

Au total, les personnes enquêtées disposent de 35 lexèmes pour nommer les types d'amandiers à l'échelle du pays. Nous avons regroupé sous une seule dénomination les variantes phonétiques simples autour de 'Ferk' ('Ferki', 'Frika', 'Frik', 'Fark'). Lors des entretiens pour établir des listes libres, les informateurs ont d'abord cité le nom qu'ils utilisent pour chaque variété puis les synonymes qu'ils savent être utilisés ailleurs dans le pays.

Les synonymes qui renvoient aux mêmes critères de caractérisation ont été regroupés afin de réduire le biais dû à la synonymie dans les analyses: on compte ainsi cinq synonymes possibles de 'Khachabi' ('Kassi', 'Klaibi', 'Barri', 'Kasr', 'Chakouchi') et deux synonymes de 'Abrach' ('Aghbar' et 'Bou_waber'). Les termes 'Baladi' et 'Halwani', synonymes respectifs de 'Khachabi' et de 'Ferk', n'ont pas été regroupés car ces termes renvoient à différentes classes de critères de classification.

³ Ce prix peut varier selon les années : les 200g étaient vendus 3 euros en 2010 et 15 euros en Février 2018.

Tableau II-5. Résultats des listes libres à l'échelle du Liban.

| Type nommé | Fréquence | Indice de Sutrop |
|------------------|-----------|------------------|
| Ferk | 0,55 | 0,32 |
| Khachabi | 0,69 | 0,30 |
| Halwani | 0,37 | 0,25 |
| Awja | 0,40 | 0,16 |
| Itali | 0,20 | 0,08 |
| Baladi | 0,08 | 0,04 |
| Oum Omar | 0,12 | 0,03 |
| Nahali | 0,08 | 0,03 |
| Morr | 0,11 | 0,03 |
| Abrach | 0,06 | 0,02 |
| Istanbouli | 0,08 | 0,02 |
| Mdaabal | 0,05 | 0,02 |
| Laouz | 0,02 | 0,02 |
| Mallisi | 0,05 | 0,01 |
| Bandouk Awja | 0,03 | 0,01 |
| Cheikh-el-achiri | 0,03 | 0,01 |
| Faransi | 0,03 | 0,01 |
| Amercani | 0,02 | 0,01 |
| Nos-Ferki | 0,02 | 0,01 |
| Afsaji | 0,02 | 0,01 |
| Nos-Khachabi | 0,02 | 0,01 |
| Ajami | 0,02 | 0,00 |
| Oum Rachid | 0,02 | 0,00 |
| Masri | 0,02 | 0,00 |

La première colonne correspond aux noms cités ; la seconde correspond aux fréquences de noms cités par rapport au total des noms collectés ; la troisième colonne correspond à l'indice de Sutrop, le type nommé le plus important à la valeur la plus élevée.

Après regroupement, on constate que les informateurs peuvent nommer et différencier 24 types d'amandier à l'échelle du pays. L'indice de Sutrop montre que les types nommés les plus fréquents et mentionnés en premier dans les listes libres sont : 'Khachabi', 'Ferk', 'Halwani' et 'Awja'. (Tableau II-5). A l'échelle du village de Zeghrine, les 13 informateurs ne citent en revanche que 10 types nommés pour des listes d'une taille moyenne égale à 5 (variant de 3 à 7 termes) (Tableau II-6). 'Halwani' est ici le plus fréquemment cité suivi d' 'Amercani', de 'Faransi' et d' 'Itali' qui ont les mêmes fréquences, puis de 'Oum Omar' et de 'Abrach'. L'indice de Sutrop le plus élevé est pour 'Halwani' suivi de 'Faransi', 'Itali', 'Amercani' et 'Oum Omar' (Figure II-2).

1) Les types nommés par zone géographique

Les 19 informateurs localisés dans la région du Mont Liban citent pour leur part 16 types nommés pour des listes d'une taille moyenne égale à 3 (variant entre 2 et 6 termes). 'Khachabi' est le plus fréquemment cité et a l'indice de Sutrop le plus élevé ; suivi de 'Ferk' et de 'Awja' (Figure II-3a).

Les 18 informateurs de la région du Liban nord citent eux aussi 16 types nommés avec une taille moyenne de liste égale à 4 (variant de 1 à 7). 'Khachabi' est le plus fréquemment cité suivi de 'Halwani', d' 'Itali' et de 'Awja' qui ont les mêmes fréquences et de 'Faransi'. L'indice de Sutrop le plus élevé est pour 'Halwani' suivi de 'Ferk', 'Khachabi' et 'Itali' (Figure II-3b).

Les 15 informateurs de la région de la Bekaa citent 12 types nommés avec une taille moyenne de liste égale à 3 (variant de 1 à 4). 'Khachabi' est le plus fréquemment cité et a l'indice de Sutrop le plus élevé, suivi de 'Halwani' et 'Ferk' (Figure II-3c).

Les 12 informateurs de la région du Liban sud citent 5 types nommés avec une taille moyenne de liste égale à 2 (variant de 1 à 3). 'Ferk' est le type nommé les plus fréquemment cité et a l'indice de Sutrop le plus élevé, suivi de 'Khachabi' (Figure II-3d).

Tableau II-6. Nombre de listes libres collectées dans les différentes zones éco-géographiques du Liban et à Zeghrine.

| | Mont Liban | Liban Nord | Bekaa | Liban Sud | Total | Zeghrine |
|------------------------------|------------|------------|-------|-----------|-------|----------|
| Nombre de listes | 19 | 18 | 16 | 12 | 65 | 13 |
| Longueur moyenne des listes | 3 | 4 | 3 | 2 | 3 | 5 |
| Nombre de types nommés | 16 | 16 | 12 | 5 | 24* | 10 |
| Nombre total de types nommés | 59 | 67 | 45 | 27 | 198 | 59 |

* nombre de variétés citées sans répétition.

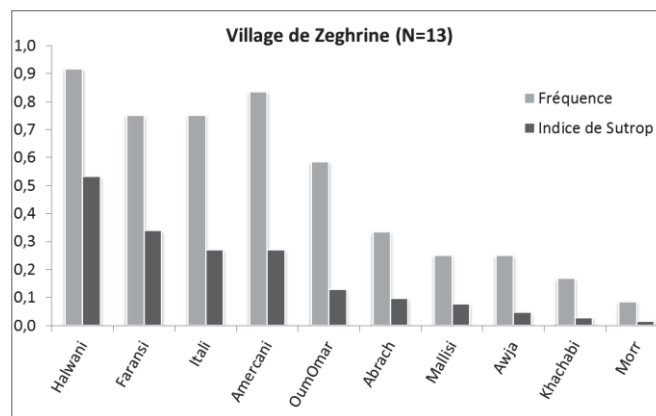


Figure II-2. Les types nommés cités à Zeghrine. Triés de gauche à droite par ordre décroissant d'indice de Sutrop.

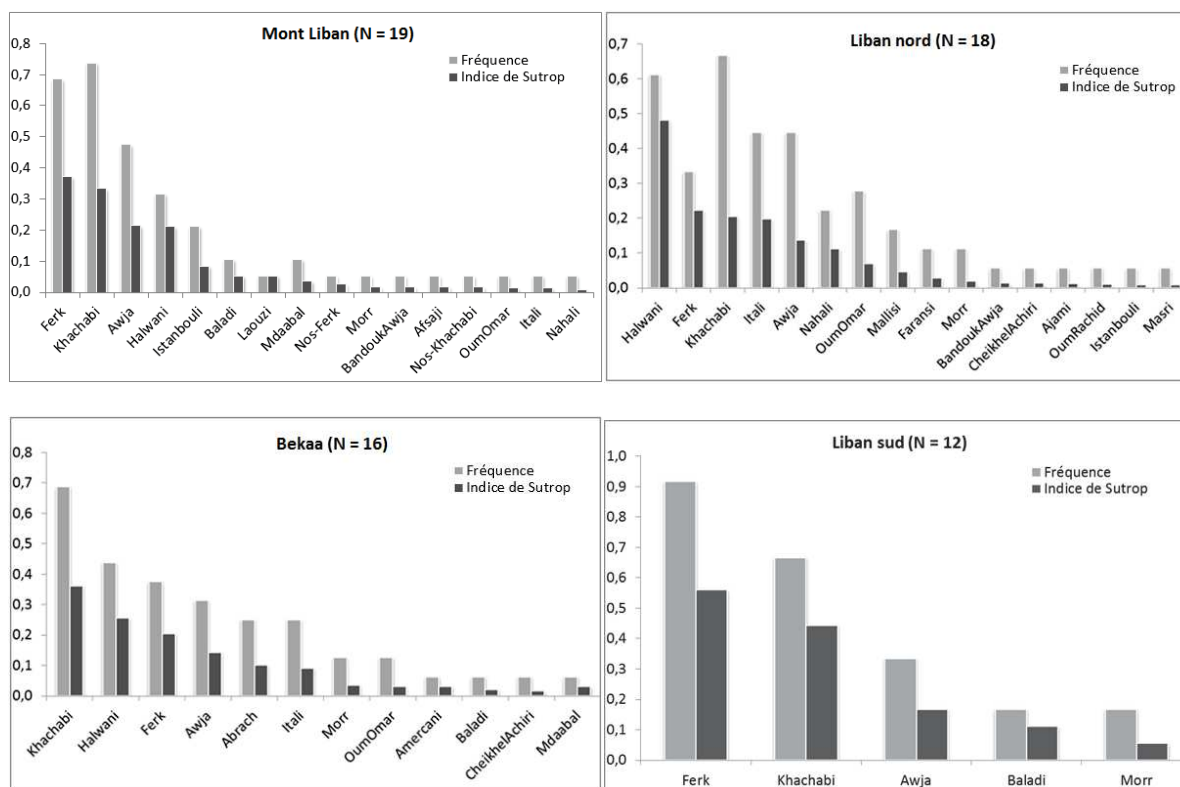


Figure II-3. Les types nommés cités par région éco-géographique. Triés de gauche à droite par ordre décroissant d'indice de Sutrop.

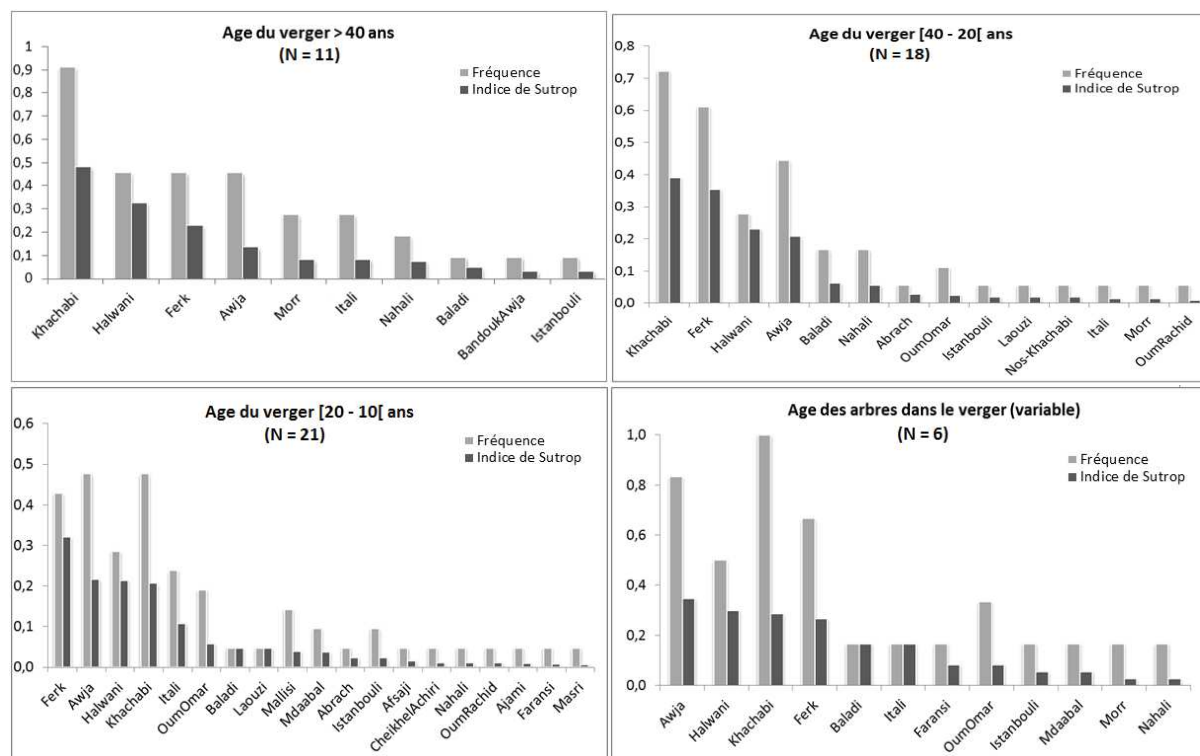


Figure II-4. Les types nommés cités par âge du verger. Triés de gauche à droite par ordre décroissant d'indice de Sutrop

2) Les types nommés par classe d'âge du verger

Les 11 informateurs, dont les vergers sont âgés citent 10 types nommés avec une taille moyenne de liste égale à 3 (variant de 1 à 6). 'Khachabi' est le plus fréquemment cité et a l'indice de Sutrop le plus élevé. 'Halwani' et 'Ferk' ont les mêmes fréquences, par contre en se basant sur l'indice de Sutrop, 'Halwani' a un rang plus élevé dans la liste, juste après 'Khachabi', suivi de 'Ferk' (Figure II-4a).

Les 18 informateurs dont les vergers ont entre 20 et 40 ans citent 14 types nommés avec une taille moyenne de liste égale à 3 (variant de 2 à 6). 'Khachabi' est le plus fréquemment cité et a l'indice de Sutrop le plus élevé ; suivi de 'Ferk'. 'Awja' est le troisième type nommé le plus fréquent ; par contre 'Halwani' est le troisième type nommé le plus cité dans les listes libres (Figure II-4b).

Les 21 informateurs dont les vergers sont récents citent 19 types nommés avec une taille moyenne de liste égale à 3 (variant de 1 à 7). 'Khachabi' et 'Awja' sont les plus fréquemment cités suivi de 'Ferk'. En se basant sur l'indice de Sutrop, 'Ferk' est le plus cité en premier dans les listes libres, suivi de 'Awja' et 'Halwani' (Figure II-4c).

Les six informateurs dont les vergers sont d'âge variable citent 12 types nommés avec une taille moyenne de liste égale à 4 (variant de 3 à 6). 'Khachabi' est le type nommé le plus fréquent suivi d'Awja', 'Halwani' et 'Oum Omar'. L'indice de Sutrop montre qu'Awja' est le plus souvent cité en premier dans les listes libres suivi de 'Halwani', 'Khachabi' et 'Ferk' (Figure II-4d).

3) Les types nommés par type de verger

Les 34 informateurs qui ont des vergers familiaux ont cité 16 types nommés avec une taille moyenne de liste égale à 3 (variant de 1 à 6). 'Khachabi' est le type nommé le plus fréquent

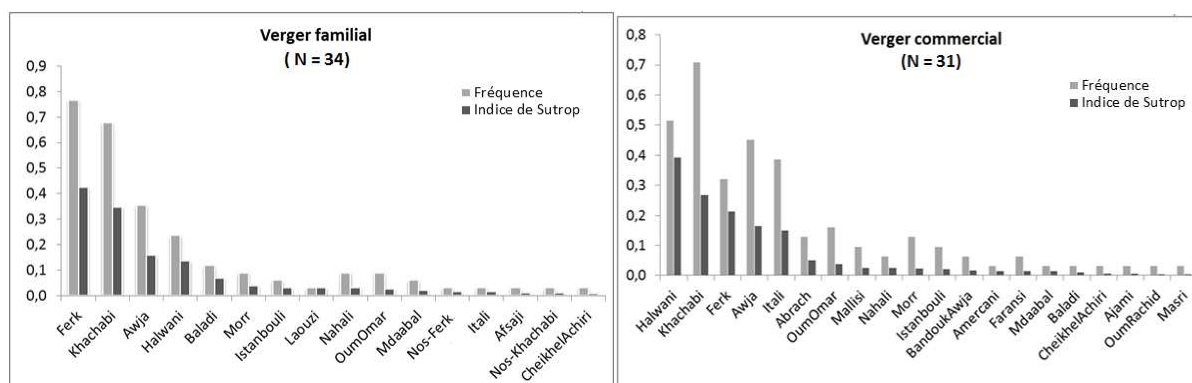


Figure II-5. Les types nommés cités dans les deux types de vergers, familial et commercial. Triés de gauche à droite par ordre décroissant d'indice de Sutrop.

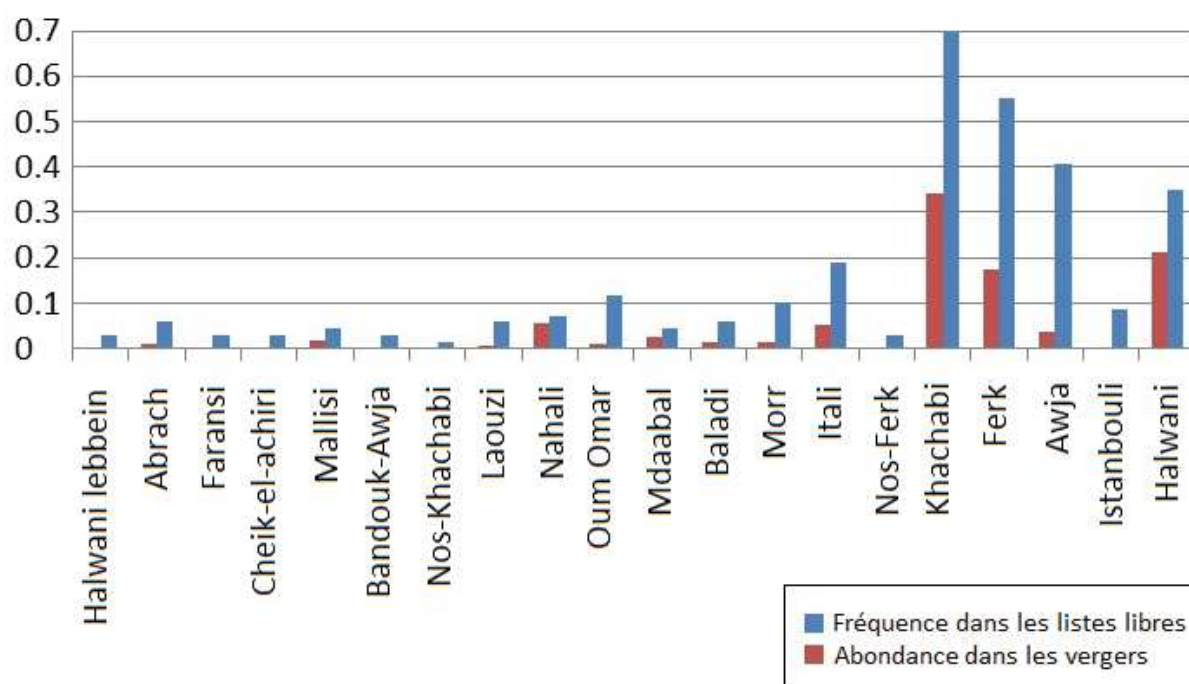


Figure II-6. Histogramme des types nommés représentant la fréquence du type nommé dans la liste libre et pour son abondance dans les vergers.

suivi de 'Halwani' et 'Awja'. L'indice de Sutrop montre que 'Halwani' est le plus cité en premier dans les listes libres suivi de 'Khachabi' et 'Ferk' (Figure II-5a).

Les 31 informateurs qui ont des vergers commerciaux ont cité 20 types nommés avec une taille moyenne de liste égale à 3 (variant de 1 à 7). 'Ferk' est le type nommé le plus fréquemment cité et a l'indice de Sutrop le plus élevé, suivi de 'Khachabi' et 'Awja' (Figure II-5b).

Les informateurs qui exploitent des vergers familiaux ont de même cité des noms qui font référence aux types nommés les plus fréquents globalement, ils ont cité en plus 'Nos-Khachabi' et 'Nos-Ferk' (dureté de coque intermédiaire entre 'Khachabi' et 'Ferk') (Figure II-5a). Les informateurs qui exploitent des vergers commerciaux sont les seuls à citer les types nommés 'Faransi', 'Amercani', 'Mallisi', 'Abrach' (Figure II-5b).

Catégories des types nommés

Les types les plus fréquemment nommés sont également les plus abondants dans les vergers (Figure II-6). Certains types nommés sont cultivés par la majorité des agriculteurs dans les différentes régions du pays : c'est le cas de 'Khachabi', 'Halwani', 'Awja', 'Ferk' ; d'autres sont cultivés dans une seule région: c'est le cas de : 'Nahali', cultivé dans les villages de Batroun au Liban nord, de 'Oum Omar', cultivé à Akkar au Liban nord, et aussi de 'Amercani', cultivé à Zeghrine dans la Bekaa. D'autres types nommés sont cultivés de manière marginale, par un seul agriculteur de notre échantillon (c'est le cas de : 'Oum Rachid', cultivé à Bakhoun au Liban Nord, et de 'Oum Hsein', cultivé a Zouk Mikael au Mont Liban). Par ailleurs, certains types nommés sont abondamment cultivés dans les vergers (cas de 'Halwani') alors que d'autres sont limités à quelques arbres par verger, toujours associés à d'autres types nommés ('Awja' par exemple).

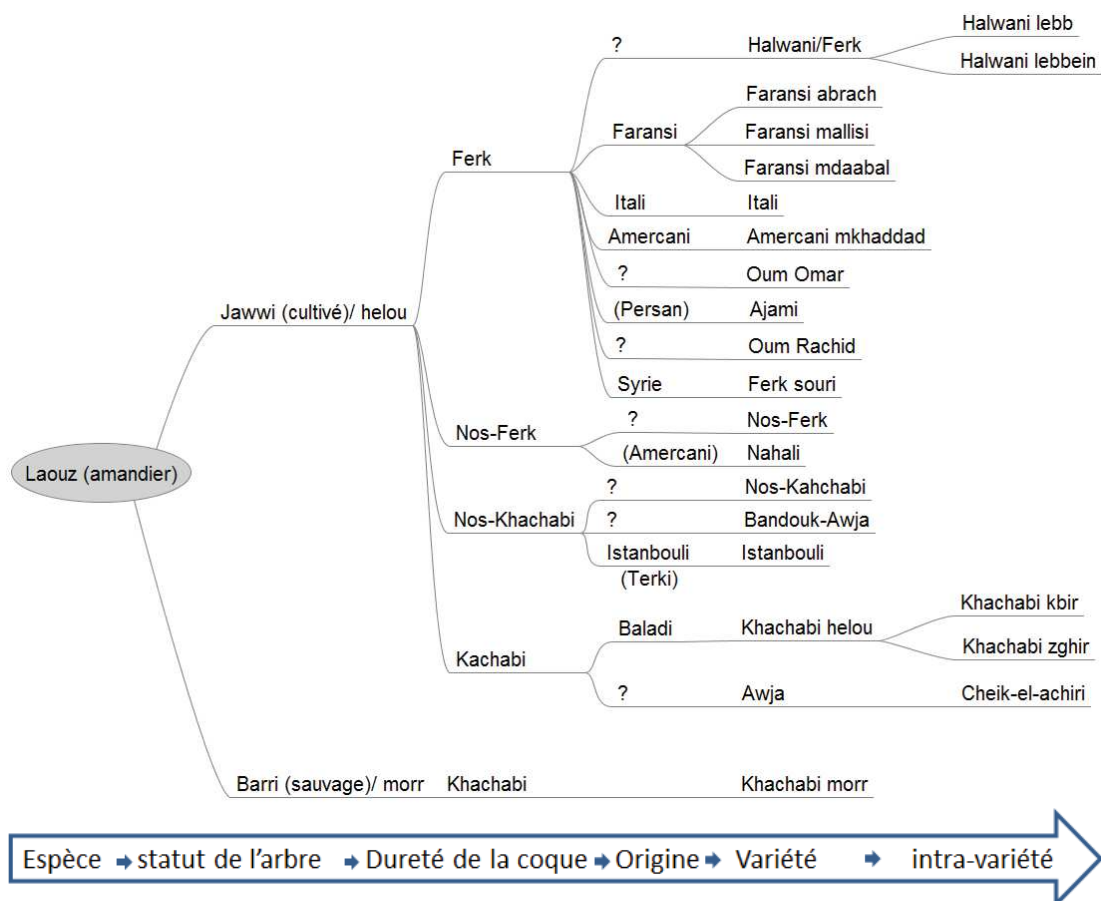


Figure II-7. Proposition de Taxinomie locale de l'amandier au Liban. Ce graphique se lit de droite à gauche. Le point d'interrogation (?) est pour les types nommés dont l'origine est inconnue.

On identifie ainsi trois grandes catégories de types nommés en fonction de leur fréquence d'occurrence dans les listes libres et de leur abondance effective dans les vergers (Figure II-6) à une échelle nationale ou locale:

- Les types nommés qui sont à la fois fréquents et abondants à l'échelle du pays (cas de : 'Halwani', 'Khachabi'). Quatre autres types nommés ('Nahali', 'Oum Omar', 'Abrach', 'Amercani') appartiennent à cette classe car ils sont fréquent et abondants mais à une échelle locale;
- Les types nommés fréquents mais peu abondants. Le type nommé 'Awja' est parmi les plus cités dans les listes libres mais sa présence est limitée à quelques arbres par verger;
- Les types nommés peu fréquents et peu abondants (cas de : 'Oum Rachid', 'Oum Hsein',...). Quatre autres types nommés dans cette dernière catégorie ('Bou Soumegh', 'Melkani', 'Biadi' et 'Metwi') sont signalés dans la littérature (Chalak *et al.* 2007) mais n'ont pas été rencontrés lors de nos enquêtes et inventaires.

Nomenclature et taxonomie locale

Nos enquêtes révèlent que la nomenclature locale de l'amandier (Figure II-7) se base sur un monolexème (ex : 'Laouz') ou sur un dilexème (ex : 'Bandouk Awja', 'Oum Omar', 'Oum Hsein', 'Cheikh-el-achiri'). Nous avons aussi constaté que les informateurs avaient tendance à donner le nom de l'espèce lorsqu'ils n'arrivaient pas à construire une liste de variétés d'amandier. Ils utilisent alors le terme de base « 'Laouz' » ou « 'Luz' » (amande) ; unité taxonomique minimale pour reconnaître l'arbre (l'amandier) en langue arabe et le distinguer d'autres espèces. Les informateurs différencient ensuite l'amandier sauvage 'Barri' de l'amandier cultivé 'Jawwi'.

Le type sauvage est reconnaissable à sa coque dure et à son amandon amer ('Khachabi morr'), il se propage par semis et s'utilise comme porte-greffe. Les variétés de l'amandier cultivé se regroupent en deux catégories englobantes : l'une à coque dure et amandon doux ('Khachabi helou' et ses synonymes), l'autre à coque tendre ('Ferk'). Ces deux catégories correspondent aussi à des types nommés: 'Khachabi', qui est locale ('Baladi'), qui se propage par semis et peut être utilisée comme porte-greffe. 'Ferk' correspond au type nommé 'Halwani', (amandon doux) qui se propage par plants greffés ou greffons.

Certains noms cités comme déterminants ont été cités aussi en terme de base, dans ce cas on a considéré que le nom cité en terme de base et le déterminant se réfèrent à une même variété (ex : 'Faransi Abrach' ou 'Abrach', la variété est Abrach et Faransi indique son origine ; de même pour 'Mkaddad' ou 'Amercani Mkhaddad', la variété est Mkhaddad et Amerani indique son origine). Par contre quand le nom est cité uniquement en déterminant, on a considéré que le terme de base correspond à la variété et que le déterminant correspond au niveau intra-variétal (ex : la variété 'Khachabi' qui est divisée en deux sous variétés 'Khachabi kbir' et 'Khachabi zghir').

Nous proposons ainsi une taxonomie locale de l'amandier au Liban (Figure II-7), présentant différents niveaux taxonomiques :

- Niveau 1: correspond à l'espèce ('Laouz').
- Niveau 2: distingue le statut de l'arbre lui-même dans la nature ou dans le verger: 'Barri' (sauvage) ou 'Jawwi' (cultivé). Ces deux statuts se distinguent aussi par un critère gustatif de l'amandon: le 'Barri' est 'morr' (amer), il est utilisé comme porte greffe ; le 'Jawwi' est 'helou' (doux), il est consommé. Ce niveau correspond à une catégorie englobante.
- Niveau 3: relatif à la dureté de la coque, critère principal de distinction entre les variétés 'Khachabi', à coque dure, et 'Ferk', à coque tendre. Pour le terme 'Khachabi' il est attribué à du

matériel 'Barri' (sauvage) et à du matériel 'Jawwi' (cultivé). Ce niveau correspond aussi à une catégorie englobante.

- Niveau 4: relatif aux origines, distingue ce qui vient d'ici (Baladi) de ce qui vient d'ailleurs (autres régions ou pays d'origine).

- Niveau 5: correspond à différentes variétés nommées : le nom 'Kachabi' se retrouve aussi à ce niveau. Le nom 'Ferk' réapparaît à ce niveau en tant que synonyme de 'Halwani'. La majorité des dialectes se retrouvent dans ce niveau (Ex: 'Oum Omar', 'Cheikh-el-achiri', 'Oum Rachir', Faransi Abrach, etc, Figure II-7).

- Niveau 6: correspond au niveau intra-variétale, le terme de base du niveau précédent est associé à un déterminant qui permet de donner plus de précision pour différencier les variétés en prenant en compte des critères cette fois morphologiques du fruit (amande) (grand / petit). Exemple : 'Khachabi kbir' / 'Khachabi zghir'.

Finalement, cette représentation taxonomique en arborescence (Figure II-7) livre différents niveaux de dénomination et de précision descriptive qui permettent de bien mettre en évidence des variétés regroupées sous des catégories plus englobantes. C'est aussi une représentation d'une sorte de gradation de l'espèce sur trois axes : du sauvage au cultivé, d'ici et d'ailleurs, de la diversité variétale et intra-variétale.

Cette complexité taxonomique est révélatrice ainsi de l'importance de l'amandier dans la société, aussi bien pour les façons de traiter les arbres que pour la diversité des échanges et formes de consommation des amandes.

Discussion

Dans cette étude nous avons identifié et analysé les modes de classification de l'amandier par l'étude des taxonomies locales au Liban nous permettant de comprendre la perception de la diversité intra-spécifique de cette plante et d'en mesurer l'importance pour la société.

Diversité des listes libres

La diversité variétale de l'amandier cultivé est importante dans trois sur quatre régions éco-géographiques. Le niveau du système de dénomination et de classification identifié dans les régions du Mont Liban, Liban nord et Bekaa développé par les agriculteurs, est associé à des valeurs socio-culturelles et économiques. Par contre, le système de dénomination et les critères de classification de l'amandier sont peu développés dans la région du Liban sud. Les informateurs citent des types nommés qui correspondent à l'espèce ainsi qu'aux variétés 'Khachabi', 'Ferk' et 'Awja' ; par contre la catégorisation au niveau intra-variétal est absente. Les critères de catégorisation correspondent le plus souvent à la morphologie (dureté de la coque, forme du fruit). La dominance de l'oléiculture au détriment de la culture des autres arbres fruitiers dans le Liban sud peut expliquer la faible importance culturelle de l'amandier dans cette région.

A l'échelle du pays, 'Khachabi' et 'Ferk' sont les plus fréquemment cités et sont aussi les premiers dans les listes libres. L'importance culturelle de 'Khachabi' et de 'Ferk' a été confirmée à l'échelle des régions, pour les vergers traditionnels comme pour les vergers commerciaux quel que soit leur âge, sauf pour 'Khachabi' chez les informateurs qui cultivent des vergers récents. 'Awja' est le troisième type nommé le plus important culturellement à l'échelle du pays et dans la région du Mont Liban. Ce type nommé a été fréquemment cité et se trouve parmi les premiers dans les listes libres des informateurs qui s'occupent des vergers commerciaux, des vergers récents ou des vergers d'âge variable.

Le développement de la culture de l'amandier au cours des dernières décennies dans la plaine de la Bekaa a conduit à une différenciation des producteurs. L'encadrement des agriculteurs par une coopérative de pépiniéristes a permis une modernisation agricole avec des vagues d'introduction de nouvelles variétés depuis les années 1990. Ceci a eu pour conséquence

l'émergence de vergers commerciaux évoluant vers une logique de diversification variétale pour assurer une production étalée.

Les types nommés 'Kachabi', 'Ferk' et 'Awja' n'ont pas été cités parmi les plus importants culturellement chez les informateurs de Zeghrine. L'étude menée à cette échelle nous a permis d'identifier l'introduction récente du type nommé 'Amercani Mkaddad' qui n'a été nommé qu'à Zeghrine. Les types nommés les plus cités et les premiers dans les listes libres sont 'Halwani' suivi de 'Faransi', 'Itali' et 'Amercani'; à part 'Halwani', la liste des types nommés est différente de celles cités à l'échelle du pays et des régions. La saillance des types nommés 'Faransi', 'Itali' et 'Amercani', est liée à la mémoire récente reliée aux usages à des fins commerciales. Sousa *et al.* (2016) montrent que les premières plantes médicinales citées par les Carão au Brésil sont celles les plus utilisées dans un passé proche; ainsi les auteurs constatent que l'ordre des items cités dans la liste libre est associé à la mémoire récente qui est reliée aux usages. La différence entre les variétés importantes dans le village de Zeghrine (où l'amandier est depuis longtemps central dans l'économie locale) et celles qui le sont dans le reste du pays confirme comment la diversité intra-spécifique nommée de l'amandier évolue avec l'importance et la durée de l'interaction entre les agriculteurs et cette culture.

'Khachabi' et 'Halwani'/'Ferk', types nommées phares à l'échelle du pays

La revue Jatba (actuellement revue d'ethnoécologie) a publié en 1921, un article sur la culture des arbres fruitiers en Syrie (actuellement Liban et Syrie) et Cilicie, l'enquête a été effectuée par les Services du général Gouraud (haut-commissaire de la république Française en Syrie et Cilicie). Les résultats montrent que la diversité variétale de l'amandier est limitée à trois types nommés : « Laouz morr à coque très dure et amande amer, variété la plus abondante consommée à l'état frais ; Laouz helou à coque tendre et amande douce, arbre peu résistant et peu cultivé; Laouz ferk à coque tendre et amande douce, variété très appréciée, consommée à l'état frais et à l'état sec et très utilisée en pâtisserie » (Services de général Gouraud 1921). En

se basant sur la dénomination, les critères de caractérisation et d'identification nous pensons que les trois types nommés identifiés à l'époque coloniale correspondent aux types nommés actuellement 'Khachabi morr', 'Halwani' et 'Ferk' qui sont les plus importants. Ainsi, l'importance culturelle de ces types nommés d'amandier cultivé peut être expliquée par leur ancienneté dans les vergers libanais.

Les trois types nommés 'Khachabi', 'Halwani' et 'Ferk' sont les plus fréquents dans les listes libres ; de plus leur importance culturelle est confirmée par leurs indices de Sutrop élevés. Par ailleurs nos relevés de la composition variétale des vergers montrent que ces variétés sont les plus cultivées à l'échelle du pays. Ainsi, 'Khachabi' et 'Halwani'/'Ferk' sont les types nommés phares de l'amandier cultivé à l'échelle du pays.

Cependant, ces types nommés n'ont pas été notés parmi les plus importants culturellement chez les informateurs de Zeghrine. A Zeghrine, 'Halwani' a gardé sa place citée le premier dans les listes libres car ce type nommé est très prisé sur le marché local. D'autres types nommés à fort intérêt commercial ont été cités avant 'Khachabi' qui n'est pas destiné à la commercialisation.

1) Le type nommé 'Khachabi'

Les nombreux synonymes de 'Khachabi' s'expliquent par la présence ancienne et abondante de ce type nommé partout au Liban. Pour installer ou renouveler leurs vergers les agriculteurs s'approvisionnent en graines dans leurs anciens vergers et dans les populations de montagne; chaque communauté a donné un nom qui est basé sur le critère de la dureté de la coque. Haouane (2012), a aussi expliqué la présence de synonymie entre les noms donnés aux variétés d'olivier au Maroc par la diversité des groupes sociaux.

'Khachabi' consommé vert et sec, est maintenu grâce à la propagation par semis dans les vergers familiaux traditionnels. 'Khachabi' est présent dans certains vergers commerciaux

d'une part pour des raisons sociales (affectives) ou agronomiques (pollinisation), d'autre part, 'Khachabi' est utilisé comme porte-greffe bien adapté aux conditions phytopathologiques et pédoclimatiques du pays. Ce type nommé présente une diversité intra-variétale sur les critères de la taille du fruit.

2) Le type nommé 'Halwani'

La différence entre 'Laouz helou' et 'Lawz ferk' se faisait dans le passé (Services du général Gouraud 1921), par contre ces deux types nommés équivalents à 'Halwani' et 'Ferk' sont considérés aujourd'hui comme synonymes par les informateurs eux-mêmes. Le croisement avec une approche génétique serait nécessaire pour identifier si ces deux types nommés constituent des entités génétiques différentes. 'Halwani' est la première variété introduite à coque tendre, largement cultivée et propagée par greffage. Très appréciée par les Libanais, 'Halwani' se consomme vert, frais et sec : le fruit (amande) a un goût acidulé quand il est consommé vert. Une fois la coque durcie, l'amandon se consomme frais ou le fruit est conservé sec et se consomme toute l'année. L'amandon est utilisé partout dans le pays comme décoration sur des plats et desserts divers (Chalak & Hamadeh 2015).

Taxonomie locale de l'amandier

La taxonomie locale proposée représente la diversité de la nomenclature de l'amandier présent au Liban. Cette nomenclature est basée sur des critères de classification flexibles entre les différentes catégories. Les informateurs différencient deux types d'amandier, le cultivé et le sauvage. La nomenclature de l'amandier sauvage est motivée mais peu profonde. Par contre, l'amandier cultivé présente la diversité intra-spécifique et un système de nomenclature complexe. La flexibilité reflète la diversité et la variabilité des connaissances des informateurs. La flexibilité se trouve au niveau de 'Khachabi' et 'Ferk', qui sont des variétés mais aussi des catégories englobantes. Le nom 'Khachabi' est donné au matériel local, qu'il soit cultivé dans des vergers ou présent dans la nature (considéré comme sauvage ou feral).

Cette flexibilité peut s'expliquer aussi par le fait que les effets de la domestication sur l'amandier sont diffus et que la douceur de l'amande est le trait crucial de domestication chez cet arbre (Zohary & Spiegel-Roy 1975).

1) Catégories englobantes et dichotomie

Selon Lévi-Strauss (1962), la classification vernaculaire est souvent basée sur des oppositions simples. Dans notre travail, les informateurs ont cité des oppositions pour certains critères de catégorisation. La première opposition est biologique, à ce niveau s'opposent 'morr' (amer) / 'helou' (doux) ; au même niveau, 'Barri' / 'Jawwi' sont utilisés pour différencier l'amandier sauvage amer de l'amandier cultivé doux.

'Khachabi' / 'Ferk', ces critères englobants qui différencient à l'intérieur du cultivé en se basant sur le critère « dureté de la coque » sont communs aux différentes variétés de l'amandier, et utilisés aussi pour deux taxons au niveau variétal de la taxonomie locale.

Baladi/ autres origines: en opposition à ce qui vient d'ici, du pays, les informateurs ont souvent cité l'origine du matériel qui vient d'ailleurs. 'Baladi' (du pays) est un terme employé en arabe (dialecte libanais) pour nommer tout ce qui est local. En opposition, les informateurs ont cité l'origine pour différencier ce qui provient d'ailleurs : de la France 'Faransi', de l'Italie 'Itali'/'Telyani' et de l'Amérique 'Amercani'. La dénomination suivant l'origine confirme l'introduction récente de ces types nommés dans l'ensemble des variétés d'amandier cultivés au Liban. Simenel (2010) a noté une opposition similaire que les Marocains utilisent pour différencier ce qui vient d'ici et ce qui vient d'ailleurs. Ainsi, «beldi» est employé fréquemment dans la langue tachelhit mais aussi dans l'arabe dialectal du Maroc pour désigner dans le monde végétal ou animal «ce qui vient du terroir» en le différenciant littéralement de «roumi» pour qualifier des produits d'origine étrangère «ce qui vient de l'Occident» (Simenel 2010). Cette opposition a été validée pour l'amandier au Maroc, la

société Bni Boufrah nomme l'amandier local « bledi » tandis que tout ce qui vient d'ailleurs est « rouni » (Delplancke & Aumeeruddy-thomas 2017). Pour le type nommé 'Istanbuli', le nom montre que ce matériel a été introduit de la Turquie pendant l'occupation ottomane. A l'époque où tous les ordres militaires venait d'Istanbul, les libanais attribuaient le nom Istanbul à tout ce qui venait de la Turquie.

2) Evolution de la nomenclature

La classification de l'amandier est complexe et se base sur des termes motivés de trois catégories de critères 1) morphologiques et organoleptiques, 2) d'origine et 3) métaphore, utilisées pour différencier les taxons de niveau variétal. D'autres catégories de critères sont utilisées pour l'identification, et ne figurent pas dans la nomenclature telle que l'usage et l'importance économique.

Les critères morphologiques et organoleptiques sont les plus utilisés pour nommer les types les plus anciens; les informateurs nomment le 'Khachabi' cultivé dans leur verger, 'Khachabi helou' (à amandon doux) et le 'Khachabi' présent dans la nature, 'Khachabi morr' (à amandon amer). La douceur de l'amande est le principal critère lié à la domestication de l'amandier, (Grasselly & Crossa-Raynaud 1980). Un autre critère de classification important est la dureté de la coque ce critère est aussi biologique, il est parmi les critères descripteurs morphologiques de l'IPGRI (Gülcan 1985). La caractérisation de la diversité morphologique de 36 individus d'amandier cultivé au Liban a permis d'identifier une diversité de la dureté de la coque qui varie de l'extrêmement dure, dure, intermédiaire, à tendre (Chalak *et al.* 2007).

Critère d'origine et adoption :

Le système de nomenclature a évolué, les variétés introduites sont souvent nommées par le nom du pays d'origine ('Faransi', 'Itali', 'Amercani', etc.). Avec l'introduction de plusieurs variétés provenant d'un même pays, le processus de catégorisation a évolué, les agriculteurs

ont ajouté un caractère d'identification morphologique, « un déterminant » pour différencier les variétés introduites du même pays ('Faransi Abrach' et 'Faransi Mallisi'). Cependant, certains informateurs ont nommé la variété introduite par le déterminant uniquement, ils n'ont pas donné le terme de base qui correspond à l'origine. Néanmoins, le nom du pays d'origine a priori a été mentionné lorsqu'on a demandé des détails sur le type nommé. C'est le cas de 'Faransi Abrach' qui a été nommé 'Abrach' par deux informateurs qui ont cité la variété dans la liste libre. Suite à ces résultats, nous avons fait l'hypothèse que cette catégorie de types nommés est récente. On suppose que l'élimination du nom du pays d'origine de la dénomination peut parvenir en deuxième étape du processus d'adoption du matériel végétal (comme pour 'Halwani', introduite à origine inconnu). Ainsi, l'adoption de la variété introduite d'Amérique au début du siècle dernier (selon l'informateur de Kfarhatna) a mené à une appropriation par l'attribution d'un nom local, 'Nahali', qui correspond au nom de la personne qui l'a introduite.

Pour une identification des sélections locales

Le système de nomenclature contient des types nommés qui peuvent refléter des sélections locales. Selon les agriculteurs le type nommé 'Bandouk Awja' est un hybride entre 'Awja' et un autre type nommé. Selon l'agriculteur de Halat dans la région de Mont Liban : « J'associe 'Bandouk Awja' et 'Halwani' dans le même verger pour une meilleure production de 'Halwani' ; Je pense que 'Bandouk Awja' est un descendant de 'Awja' et 'Halwani' ». Un autre exemple est celui du type nommé 'Cheikh-el-achiri' ; selon l'agricultrice de Wadi el Rihan dans la région du Liban Nord: « Un agriculteur de la région a trouvé du 'Awja kbir' (grand) dans son verger, il l'a ensuite propagé et les gens ont nommé cette nouvelle variété 'Cheikh-el-achiri' (chef du clan) ». De plus, la sélection locale de matériel apprécié est identifiée chez un agriculteur du village de Sfarai dans la région du Liban Sud qui ne cultive que du 'Halwani lebbein' (à amandon double).

Conclusion

Cette étude nous a permis d'évaluer l'importance culturelle de l'amandier cultivé au Liban, et d'identifier sa diversité intra-spécifique telle qu'elle est perçue par les informateurs qui sont en majorité des agriculteurs. Les informateurs de Zeghrine citent en moyenne 5 types nommés, tous d'intérêt économique. Ces informateurs s'occupent de vergers commerciaux et sont dans une dynamique de diversification. Les autres groupes d'informateurs citent cependant en moyenne seulement 3 types nommés. Les types nommés 'Khachabi' et 'Ferk' / 'Halwani' sont les plus cités et nous pouvons les considérer comme des types phares. Notre démarche a en effet permis à la fois d'évaluer la diversité de l'amandier cultivé au Liban et de mesurer la fréquence et l'abondance des types nommés recensés. Ce résultat vient enrichir les travaux antérieurs à cette recherche.

L'hétérogénéité des connaissances des informateurs a conduit à proposer une taxonomie locale flexible. La flexibilité de la taxonomie locale est révélée par la présence de catégories englobantes et par la complexité du système de nomenclature. Cette nomenclature locale se base sur des critères de catégorisation en lien avec des critères morphologiques ou organoleptiques, des caractéristiques d'origine et des métaphores. Bien que l'usage, l'intérêt économique, le mode de propagation et les caractéristiques agronomiques ne soient pas énoncés directement dans la nomenclature, ces caractéristiques correspondent à des critères d'identifications importants pour le choix du type nommé à cultiver, surtout dans les vergers commerciaux.

La taxonomie locale de l'amandier au Liban est profonde, elle met en évidence une diversité variétale importante et une diversité intra-variétale pour certains types nommés. Nos résultats montrent aussi que la taxonomie locale a évolué; elle se base sur les critères en lien avec la domestication et évolue en fonction de l'introduction de variétés étrangères dans le pays.

Déclaration d'éthique

Cette étude a été menée en collaboration entre le CIRAD, l'Université de Montpellier et l'Institut de Recherches Agronomiques Libanais (Ministère de l'Agriculture du Liban). L'IRAL a un mandat national pour collecter et conserver les ressources génétiques. L'étude a été menée dans le cadre du travail de B. Hamadeh à l'IRAL.

Avant de mener des entretiens, un accord verbal du maire de chaque commune prospectée a été obtenu. Sur le terrain, les agriculteurs ont été informés du cadre de cette étude par l'Institut de Recherches Agronomiques Libanais, et leur accord individuel obtenu. Aucune données personnalisées n'ont été maintenues dans les données présentées.

Remerciements

Les auteurs souhaitent remercier d'abord les informateurs qui ont participé à cette étude et qui ont partagé une partie de leurs connaissances. Nous remercions aussi les personnels administratifs des différentes mairies qui nous ont permis d'identifier certains informateurs clef. Enfin nos remerciements vont également aux membres scientifiques et logistiques de l'IRAL qui ont facilité la réalisation des enquêtes sur le terrain. Les auteurs tiennent à remercier Prof. Amis Chalak pour ses corrections des versions préliminaires du manuscrit. Merci aussi aux étudiants de l'Université Libanaise, Faculté d'Agronomie, A. Hamadeh et M. Abdallah, pour leurs contributions au terrain effectué à Zeghrine.

Bibliographie

Barnaud A., Deu M., Garine E., Chanterreau J., Bolteu J., Ouin Koïda E., Mckey D. & Joly H.I. 2009 – A Weed-Crop Complex in Sorghum: The Dynamics of Genetic Diversity in a Traditional Farming System. *American Journal of Botany* 96 (10): 1869–79. DOI:10.3732/ajb.0800284.

Barnaud A., Deu M., Garine E., Chanterreau J., Bolteu J., Mckey D. & Joly H.I. 2007 – Local Genetic Diversity of Sorghum in a Village in Northern Cameroon: Structure and Dynamics of Landraces. *Theoretical and Applied Genetics* 114 (2): 237–48. DOI:10.1007/s00122-006-0426-8.

Berlin B. 1992 – Ethnobiological Classification: Principles of Categorization of Plants and Animals in Traditional Societies. Princeton (à voir!!!), Princeton University Press, xvii + 335 p.

Brush S.B. & Perales H.R. 2007 – A Maize Landscape: Ethnicity and Agro-Biodiversity in Chiapas Mexico. *Agriculture, Ecosystems and Environment* 121 (3): 211–21. DOI:10.1016/j.agee.2006.12.018.

Brush S.B. 2004 – Farmers' Bounty. Locating Crop Diversity in the Contemporary World. Yale, Yale University Press, 352 p.

Caillon S., Quero-Garcia J., Lescure J.P. & Lebot V. 2006 – Nature of Taro (*Colocasia Esculenta* (L.) Schott) Genetic Diversity Prevalent in a Pacific Ocean Island, Vanua Lava, Vanuatu. *Genetic Resources and Crop Evolution* 53 (6): 1273–89. DOI:10.1007/s10722-005-3877-x.

Chalak L., Chehade A., Kadri A., Cosson P., Zanetto A., Dirlewanger E. & Laigret F. 2006 – Preliminary Characterization of Cultivated Almonds (*Prunus Dulcis* L.) in Lebanon by Morphological Traits and Microsatellite Markers. *Biologia Tunisie* 4: 0.

Chalak L. & Hamadeh B. 2015 – Almond Wild Relatives in Lebanon: Distribution, Uses and Main Threats. *Acta Horticulturae* 1074: 43-48. DOI: 10.17660/ActaHortic.2015.1074.5

Chalak L., Chehade A. & Kadri A. 2007 – Morphological Characterization of Cultivated Almonds in Lebanon. *Fruits* 62 (3): 177–86.

Cotton, CM. 1996 – Ethnobotany: Principles and Applications. Baffins Lane, John Wiley and Sons, 434 p.

Delplancke M., Alvarez N., Espíndola A., Joly H.I., Benoit L., Brouck E. & Arrigo N. 2012 – Gene Flow among Wild and Domesticated Almond Species: Insights from Chloroplast and Nuclear Markers. *Evolutionary Applications* 5 (4): 317–29. DOI: 10.1111/j.1752-4571.2011.00223.x

Delplancke M. & Aumeeruddy-Thomas Y. 2017 – Des Semis et Des Clones. *Revue d'ethnoécologie*: 0–27.

Elias M., Mckey D., Panaud O., Anstett M.C. & Robert T. 2001 – Traditional Management of Cassava Morphological and Genetic Diversity by the Makushi Amerindians (Guyana, South America): Perspectives for on-Farm Conservation of Crop Genetic Resources. *Euphytica* 120 (1): 143–57. DOI:10.1023/A:1017501017031.

Friedberg C. 1991 – Operative Aspects of Folk Classification. In *Man and a half: essays in Pacific anthropology and ethnobiology in honour of Ralph Bulmer*. Auckland: *The Polynesian Society*, ed. A. Pawley, 102-109.

Gouraud, Services du général. 1921 – La Culture Des Arbres Fruitiers En Syrie et Cilicie. *Revue de Botanique Appliquée et d'agriculture Coloniale* 1: 129–48. DOI:<https://doi.org/10.3406/jatba.1921.4001>.

Grasselly C. & Crossa-Raynaud P. 1980 – L'amandier: Techniques Agricoles et Productions Méditerranéennes. Paris, G.-P. Maisonneuve et Larose, 446 p.

Gülcan R. 1985 – Descriptor List for Almond (*Prunus Amygdalus*) (Revised). <http://agris.fao.org/agris-search/search.do?recordID=XF2016038642>.

Haouane H. 2012 – Origines , Domestication et Diversification Variétale Chez l'Olivier (*Olea Europaea* L.) à l'Ouest de La Méditerranée. Doctorat en Evolution – Ecologie – Ressources Génétiques – Paléontologie, ED SIBAGHE, Université de Montpellier.

Hmimsa Y., Aumeeruddy-Thomas Y., & Ater M. 2012 – Vernacular Taxonomy, Classification and Varietal Diversity of Fig (*Ficus Carica* L.) among Jbala Cultivators in Northern Morocco. *Human Ecology* 40 (2): 301–13. DOI:10.1007/s10745-012-9471-x.

Levi-Strauss C. 1962 – Lévi-Strauss : La Pensée Sauvage. Paris, Pion, 351 p.

Martin G. J. 2004 – Ethnobotany : A Methods Manual. Boston, Springer. 296 p.

Mekbib F. 2007 – Infra-Specific Folk Taxonomy in Sorghum (*Sorghum Bicolor* (L.) Moench) in Ethiopia : Folk Nomenclature , Classification , and Criteria. *Journal of Ethnobiology and Ethnomedicine* 18: 1–18. DOI:10.1186/1746-4269-3-38.

Mouterde P. 1966 – Nouvelle Flore Du Liban et de La Syrie. Volume 1. Beyrouth, Imprimerie catholique, 563 p.

Pennec F., Wencelius J., Garine E., Raimond C. & Bohbot H. 2012 – FLAME: Free-list analysis under Microsoft Excel. Paris:CNRS. <http://www.mae.u-paris10.fr/lesc/spip.php?article63>

Perales H.R., Benz B.F. & Brush S.B. 2005 – Maize Diversity and Ethnolinguistic Diversity in Chiapas , Mexico SCIENCES. *Proceedings of the National Academy of Sciences* 102: 949–

54.

Post G.E. 1932 – Flora of Syria, Palestine, and Sinai. *Volume 1 of American University of Beirut, Natural science series*, 639 p.

Simenel R. 2010 – Beldi/Roumi, Une Conception Marocaine Du Produit Du Terroir, l'exemple Des Aït Ba'amram de La Région de Sidi Ifni. *Hespéris-Tamuda* 45: 167–76.

Sousa D.C.P.d., Soldati G.T., Monteiro J.M., Araújo T.A.d.S., Albuquerque U.P. 2016 – Information Retrieval during Free Listing Is Biased by Memory: Evidence from Medicinal Plants. *PLoS ONE* 11 (11): 1–15. DOI:10.1371/journal.pone.0165838.

Sutrop U. 2001 – List Task and a Cognitive Salience Index. *Field Methods* 13 (3): 263–76. DOI:10.1177/1525822X0101300303.

Zohary D. & Spiegel-Roy P. 1975 – Beginnings of Fruit Growing in the Old World. *Science* 187 (4174): 319–27. DOI:10.1126/science.187.4174.319.

Zohary D., Hopf M. & Weiss E. 2012 – Domestication of Plants in the Old World: The Origin and Spread of Domesticated Plants in Southwest Asia, Europe, and the Mediterranean Basin. Oxford, Oxford University Press, 264 p.

Chapitre 3

Evolution of almond genetic diversity and farmer practices in Lebanon: Impacts of the diffusion of a graft-propagated cultivar in a traditional system based on seed-propagation

BMC Plant Biology (2018) 18:155

<https://doi.org/10.1186/s12870-018-1372-8>

**Bariaa Hamadeh^{1,2,3,4*}, Lamis Chalak⁵, Geo Coppens d'Eeckenbrugge⁴, Laure Benoit⁶,
Hélène I. Joly⁴**

1 Lebanese Agricultural Research Institute, Fanar, Lebanon

2 Université Montpellier 2, UMR CEFE, 34293 Montpellier cedex, France

3 Lebanese University, Doctoral School of Sciences and Technologies, Hadath, Lebanon

4 CIRAD, UMR AGAP, F-34398 Montpellier, France

AGAP, Univ Montpellier, CIRAD, INRA, Montpellier SupAgro, Montpellier, France

5 Lebanese University, Faculty of Agriculture, Dekwaneh, Lebanon

6 CIRAD, UMR CBGP, F-34398 Montpellier, France

CBGP, Univ Montpellier, CIRAD, INRA, IRD, Montpellier SupAgro, Montpellier, France

E-mail: bariaa.hamadeh@gmail.com; lamis.chalak@gmail.com; geo.coppens@cirad.fr;
laure.benoit@cirad.fr; helene.joly@cirad.fr

* [Correspondence: bariaa.hamadeh@gmail.com](mailto:bariaa.hamadeh@gmail.com)

⁴ Actual address: CIRAD, UMR AGAP, Dynamiques de la diversité, sociétés et environnements (DDSE), TA A-61/[03](#) Avenue Agropolis, F-34398 Montpellier Cedex 5, France.

1 Abstract

2 Background

3 Under cultivation, many outcrossing fruit tree species have switched from sexual reproduction
4 to vegetative propagation. Traditional production systems have persisted, where cultivar
5 propagation is based on a mixed reproductive system. For millenia, almond, *Prunus dulcis*,
6 has been propagated by seeds. Almond grafting remained of little importance until recently.
7 In Lebanon, both sexual and clonal reproductions are used for almond propagation. We used
8 15 microsatellite markers to investigate the effect of introducing graft-propagated cultivars
9 and associated practices, on the structure of the genetic diversity among and within the two
10 main Lebanese cultivars.

11 Results

12 As expected, the sexually propagated cultivar Khachabi exhibited more genotypic and genetic
13 diversity than the vegetatively propagated cultivar Halwani. It also exhibited lower
14 differentiation among populations. The distribution of clones showed that propagation modes
15 were not exclusive: farmers have introduced clonal propagation in the seed-propagated
16 cultivar while they have maintained a diversity of genotypes within populations that were
17 mostly graft-propagated. These practices are also important to avoid mate limitations that
18 hamper fruit production in a self-incompatible species.

19 ‘Khachabi’ is structured into two gene pools separated by the Lebanese mountains. As to
20 ‘Halwani’, two different gene pools were introduced. The most ancient one shares the same
21 geographic range as ‘Khachabi’; longtime coexistence and sexual reproduction have resulted
22 in admixture with ‘Khachabi’. In contrast, the more recent introduction of the second gene
23 pool in the Bekaa region followed an evolution towards more extensive clonal propagation of

‘Halwani’ limiting hybridizations. Furthermore, some pairs of geographically distant ‘Halwani’ orchards, exhibited low genetic distances, suggesting that a network of exchanges between farmers was effective on a large scale and/or that farmers brought clonal plant material from a common source.

Conclusions

Almond diversification in Lebanon is clearly related to the evolution of propagation practices adapted to self-incompatible cultivars. The comparison between both cultivars demonstrated the genetic effects of the introduction of a new cultivar and the associated grafting propagation practices. Our study provided information to develop a strategy for *in situ* conservation of cultivars and to limit gene flow from introduced material to ancient orchards.

Keywords

Population genetics, *Prunus*, agricultural practices, almond, vegetative propagation, grafting, genetic differentiation.

Background

Population genetic structure of a crop plant species is strongly determined by their mating system and mode of dissemination. Allogamous seed reproduction maintains the crop evolutive potential because it produces genetically variable progenies able to cope with and adapt to changing environments [1, 2, 3]. Characteristically, natural populations of outcrossing species present high within-population diversity and low between-population differentiation for nuclear genes [4]. Long-lived perennials uphold significantly higher levels of genetic diversity within populations than annual plants [5, 6]. In a plant population, strict or partial clonal propagation results in a non-random distribution of genotypes, i.e., genetic substructuring [7]. High levels of clonal propagation lead to reduced genotypic diversity concomitant with an increase in population differentiation and heterozygosity [7, 8].

Under cultivation, many outcrossing fruit tree species have switched from sexual reproduction to vegetative propagation [9], and the cost of the resulting clonality is precisely an absence or reduction of sexual recombination [10]. Vegetative propagation, by way of grafts or cuttings, ensures that selected traits are reliably passed true-to-type onto the next generation [11]. This, together with large-scale cultivation of genetically uniform cultivars, has increased genetic vulnerability in intensive fruit production. However, less intensive production systems have persisted, where cultivar propagation is based on a mixed reproductive system, with two interlinked compartments corresponding respectively to sexually and clonally propagated plants. In the long run, farmer management of the aforementioned compartments will determine the evolution and the population genetic structure of the species. Investigating the dynamics of such mixed sexual / clonal systems on spatial and temporal scales helps clarify these long-term processes. However, although the mating system of most clonally propagated crops is documented, their sexual reproductive ecologies are poorly known [11].

The relationship between sexual recombination and genetic diversity works in both directions. While the suppression of sexual recombination reduces genotypic diversity, mate limitations conversely affect the potential for sexual recombination in monoclonal or polyclonal populations of obligate out-crossers. When the target of cultivation is a fruit, production can be ensured if the reproductive constraints are alleviated, for example by parthenocarpy, a suppression of the self-incompatibility system, or a shift to hermaphroditism in dioecious species [10, 11]. Nevertheless, in many fruit tree species, even economically important ones whose cultivars have been propagated for centuries (e.g. pear, apple, cherry), efficient outcrossing is necessary, and sexual or clonal cultivar reproduction practices do not only concern propagation but also production potential. In modern, graft-established, horticulture, fruit production is ensured by establishing biclonal or polyclonal orchards, or inter-planting a few trees from a compatible pollen-donor. Synchronous flowering, pollinator activity control and separate harvesting of the different cultivars are then further constraints. How this problem is managed in traditional horticultural systems has largely escaped the attention of crop evolutionists.

The eastern Mediterranean Basin is documented to be one of the major centers of origin for agriculture, with the first domestication events of cereals and fruit trees occurring in the Levant about 12,000 years B.P. [12]. Almond, *Prunus dulcis* (Miller) D.A. Webb (syn. *Amygdalus communis* L. and *P. amygdalus* Batsch.) [13], is considered to be one of the earliest domesticated nut-producing trees. It is a small deciduous self-incompatible tree, native to the Mediterranean region, where *P. dulcis* and other *Prunus* species still grow in the wild [13]. For millennia, almond trees have been propagated by seeds [14]; thus, cultivated populations were domesticated a long time prior to the development of grafting [15]. Almond tree grafting was already known by the first century AD. However, it was of little importance until the beginning of the 19th century, when grafting became the propagation method used to

multiply selected genotypes [16]. Thus, while both sexual and clonal reproductions have been used for almond propagation in the Mediterranean region, the introduction of new improved and clonal cultivars in the form of grafts has become more and more frequent [16, 17]. However, little attention has been paid to the effect of these introductions and the associated grafting practices, on the structure of the genetic diversity both among and within cultivars.

Molecular tools such as nuclear microsatellite SSR (simple sequence repeat) analysis have been successfully used to assess genetic diversity as well as to reveal synonymous and erroneous labeling in cultivated *Prunus* fruit germplasm [18, 19, 20, 21, 22] and in other Mediterranean cultivated fruit species such as olive [23, 24] and fig [25]. To our knowledge, the differentiation among graft- or seed-propagated populations was only noted in a Tunisian apricot study [21]; its authors clearly distinguished two gene pools that are directly related to the different propagation modes of the Tunisian apricot cultivars. This study confirms the need to take into account the propagation mode when studying the structure and dynamics of genetic diversity. However, the genetic diversity and structure within ancient seed propagated and recent clonally propagated almond cultivars has not been compared so far. Furthermore, to our knowledge, previous studies have been carried out at the individual level and not at the population level, except for Delplancke et al [26] and Wang et al [27]. In these studies, genotyping was used to characterize clonally and sexually propagated cultivars, identified by the genotype of a single individual. In our study we set out to evaluate the evolutionary process at the population level and carried out our sampling strategy accordingly.

Spontaneous populations of *P. dulcis* and its close wild relatives, *P. orientalis* (Mill.) Koehne and *P. spartioides* (Spach) C.K. Schneid., are still found in Lebanon [28, 29, 30]. Cultivated *P. dulcis* is present in different agro-climatic areas and is part of the Lebanese landscape. Old almond trees are still found over a wide range of elevations (0 to 1600 m). So far, studies on

the diversity of almond in Lebanon have only concerned morphological traits [31] and the identification of almond cultivars [18]. These studies give a preliminary view of the wide distribution of some cultivars across the country and indicate greater varietal diversity in home gardens than in modern orchards. Using nuclear and chloroplastic SSR markers, gene flow were identified in both directions between cultivated almond species and *P. orientalis*, present in Lebanon, Syria and Turkey [26]. Moreover, the Levant was identified as a center of diversification for cultivated *P. dulcis*, since it is an important reservoir of almond genetic diversity [32].

Almond has not only been cultivated under widely divergent ecological conditions, but also the history of almond cultivation in the country is long-standing. Two main horticultural forms are recognized by farmers, based on fruit shell hardness. The most common and oldest cultivar is known as ‘Khachabi’, referring to its fruit shell ‘as hard as wood’; for centuries, it has been cultivated and reproduced by seeds, maintaining a high level of morphological variability. Since the Ottoman period (16th to early 20th century), many cultivars have been introduced to the country from different origins. Among them, the cultivar now called ‘Halwani’, was probably introduced to Lebanon during the last century; it has been widely accepted and mostly propagated by grafts, mainly-in mono-varietal populations [33].

Questions arise regarding the genetic diversity of ‘Khachabi’ (common hard shell cultivar) and ‘Halwani’ (introduced soft shell cultivar), as well as their relationships and internal structure, taking into account the influence of their respective propagation modes. Nowadays, sexual reproduction is still important in almond populations and this importance depends largely on whether farmers incorporate sexual progeny into their stocks of grafting material. This practice is common in many traditional farming systems, and has probably been continued from the origin of domestication up to our time. For the present study, we sampled

Table 1 Populations collected throughout Lebanon.

n, Number of collected individuals.

| Population name | Cultivar | Eco-geographical zone | Code | n | NDD | EDD | Age of plantation (years) |
|---------------------|----------|-----------------------|-------|----|---------|---------|---------------------------|
| Bire | Halwani | Bekaa | HB_Bi | 23 | 34.5872 | 36.2391 | 25 |
| Ferzol | Halwani | Bekaa | HB_Fe | 33 | 33.8884 | 35.9396 | 25 |
| Nabi Ayla | Halwani | Bekaa | HB_NA | 30 | 33.8819 | 35.9569 | 25 |
| Zighrine | Halwani | Bekaa | HB_Ze | 16 | 34.4261 | 36.3528 | 25 |
| Amchit | Halwani | Mount Lebanon | HM_Am | 35 | 34.1475 | 35.6464 | 48 |
| Maasriti | Halwani | Mount Lebanon | HM_Ma | 30 | 33.7483 | 35.6347 | 30 |
| Jeita | Halwani | Mount Lebanon | HM_Je | 30 | 33.9533 | 35.6478 | 70 |
| Bchannine | Halwani | North Lebanon | HN_Bc | 35 | 34.3483 | 35.8859 | 20 |
| Btaaboura | Halwani | North Lebanon | HN_Bt | 25 | 34.2742 | 35.7622 | 20 |
| Deir Qanoun el Nahr | Halwani | South Lebanon | HS_DQ | 25 | 33.3000 | 35.3136 | 20 |
| Tanbourit | Halwani | South Lebanon | HS-Ta | 15 | 33.5161 | 35.4153 | 20 |
| Bakkifa | Khachabi | Bekaa | KB_Ba | 27 | 33.4933 | 35.8192 | 60 |
| Bire | Khachabi | Bekaa | KB_Bi | 20 | 33.5842 | 35.8197 | 40 |
| El Mhaidtheh | Khachabi | Bekaa | KB_EM | 30 | 33.5569 | 35.8117 | 60 |
| Irsal 2 | Khachabi | Bekaa | KB_Ir | 15 | 34.1886 | 36.3923 | 35 |
| Kamed el Lawz | Khachabi | Bekaa | KB_KL | 33 | 33.6203 | 35.8214 | 60 |
| Lucy | Khachabi | Bekaa | KB_Lu | 35 | 33.6453 | 35.8383 | 60 |
| Shaata | Khachabi | Bekaa | KB_Sh | 18 | 34.1422 | 36.2311 | 30 |
| Assia | Khachabi | North Lebanon | KN_As | 25 | 34.2189 | 35.7856 | 50 |
| El Qalamoun | Khachabi | North Lebanon | KN_EQ | 16 | 34.3872 | 35.7864 | 70 |
| Baraachit | Khachabi | South Lebanon | KS_Ba | 30 | 33.1761 | 35.4433 | 50 |
| Blida | Khachabi | South Lebanon | KS_Bl | 30 | 33.1400 | 35.5147 | 50 |
| Chebaa | Khachabi | South Lebanon | KS_Ch | 23 | 33.3475 | 35.7492 | 80 |
| Mays al Jabal | Khachabi | South Lebanon | KS_MJ | 18 | 33.1686 | 35.5242 | 37 |
| Tanbourit | Khachabi | South Lebanon | KS-Ta | 30 | 33.5161 | 35.4153 | 30 |

a large range of cultivated individuals from different agro-climatic areas in Lebanon, and analyzed them with a set of 15 SSRs selected for their reproducibility and polymorphism in almond and other *Prunus* species, with the aim of: (1) characterizing the pattern of genotypic and genetic structure of each cultivar, ‘Khachabi’ and ‘Halwani’, and (2) assessing the influence of the propagation mode on these patterns.

Methods

Survey and sampling

Field surveys were undertaken in 2009-2010 throughout Lebanon with the aim of collecting samples representative of the most common almond cultivars, ‘Khachabi’ and ‘Halwani’. We followed farmer’s determinations in the assignation of sampled trees to a given cultivar; nomenclature was usually based on shell hardness. Farmers usually propagated hard shell almonds from seed and soft ones by grafting. According to farmers, out of 331 ‘Khachabi’ individuals, 301 were propagated by seeds, and the 30 individuals from the youngest population were propagated by grafting. Out of the 285 ‘Halwani’ individuals, 255 were propagated by grafting and the 30 individuals from the oldest population were propagated either by grafting or seeds. In all, 14 ‘Khachabi’ and 11 ‘Halwani’ populations were sampled, covering the four major agro-climatic zones (Table 1; see Additional file 1: Fig. S1 presents the locations of the collected populations). Each sampled ‘Khachabi’ orchard only contained ‘Khachabi’ trees but the presence of rare individuals from other almond cultivars could sometimes be observed in ‘Halwani’ orchards. From the information gathered from farmers, we realized that, in localities of traditional almond production, orchards were either still constituted of the original cultivar Khachabi, or they had been rejuvenated by grafting most trees with ‘Halwani’ scions. In other localities, ‘Halwani’ had been established directly in new

orchards. Our population approach led us to select orchards on the availability of a large enough number of individuals of a given cultivar.

Microsatellite amplification

Total genomic DNA was extracted from silica-dried leaves using Extract-N-Amp™ Plant PCR Kits (SIGMA - Aldrich, St. Louis, MO, USA) abiding by the supplier's instructions. Fifteen nuclear SSR markers previously selected for studying the genetic diversity of *Prunus* species in the Mediterranean region were used for this study: UDP96-001, UDP96-018, UDP96-003, UDP97-401, UDP98-408, UDP98-409, pchgms1, pchgms3, BPPCT017, BPPCT001, BPPCT007, BPPCT025, BPPCT036, CPSCT018, CPDCT045 [34, 35, 36, 37, 38]. We amplified these 15 SSR markers into three multiplexed PCRs, using one of the FAM, HEX or NED fluorophore-labeled primers (PE Applied Biosystems, Warrington, UK). Multiplexed PCRs were carried out with the Extract-N-Amp PCR Ready Mix (SIGMA - Aldrich, St. Louis, MO, USA) in a final volume of 10 µl, containing 5µl of a SIGMA Master Mix 2X, 0.4µl of primer mix at 5µM, 1.6 µl of ultrapure water and 1µl of template dNTPs. 3µl of PCR product was mixed with 15µl of formamide and 0.2µl of Genescan™ 500 LIZ size standard (Applied Biosystems, Foster City, USA), and GeneScan was performed with the ABI PRISM 3130 XL 16 capillary-sequencer (Applied Biosystems, Foster City, USA).

Genotyping and data organization

Allele size was read independently by two investigators using GEMEMAPPER 4.0 (Applied Biosystem, Foster City, USA). Genotyping errors were evaluated by checking the reproducibility of migrations using samples replicated on the different plates; in all, 95.1% of tested individuals gave the same allele size (i.e., the error rate was under 0.05).

Redundant genotypes were searched both within cultivars and among populations of both ‘Khachabi’ and ‘Halwani’ to identify multi-locus genotypes (MLGs) using GENCLONE 2.0 software [39]. Individuals with identical MLGs were considered as a clone. Different datasets were used for different analyses at different levels: global, cultivar and populations. Datasets were constructed using two different matrices: a matrix with a total of 615 individuals that will be referred to as the N dataset and a matrix with a single copy of redundant MLGs conserved within each population (with a total of 509 individuals) that will be referred to as the G dataset.

Data analysis

Genotypic diversity

The discriminant power of the 15 markers used to differentiate MLGs present in the sample was explored by plotting the number of loci versus the maximum number of MLGs detected for all datasets and for each cultivar; values were calculated using a re-sampling procedure (3000 permutations) implemented in GENCLONE. The relevance of marker information was described by the number of alleles (N_a) implemented in the software GENALEX 6.5.1 [40] and the polymorphic information contents [41], $PIC = 1 - \sum_i p_i^2 - \sum_{i,j} p_i^2 p_j^2$, where p_i and p_j are the frequency of the i th and j th alleles, implemented in CERVUS 3.0.7 software [42].

Genotypic richness was calculated following the corrected formula [43], $R = (G-1) / (N-1)$, where G is the number of distinct genotypes and N is the number of individuals sampled in a population, implemented in GENCLONE. The probability of observing two identical MLGs in each population (PI) was calculated using the GENALEX software.

Genotypic linkage disequilibrium (LD) between loci was calculated on the G dataset to avoid its overestimation; the coefficients of correlation (r) between loci were calculated within the G dataset, each cultivar and the distinct population subsets. Testing for genotypic linkage

disequilibrium (LD) using the log likelihood ratio statistic test (G-test) was performed with the software GENEPOP 4.2 [44, 45], available on the web (http://genepop.curtin.edu.au/genepop_op2.html). Exact *P*-values were estimated using the Markov Chain algorithm with 10000 dememorization steps, 1000 batches and 10000 iterations.

We conserved a single copy of each MLG (G dataset) to identify the numbers of different alleles between each pair of compared MLGs in each cultivar. The histograms of pairwise comparisons were constructed using simple matching of pairs of MLGs using Unweighted Arithmetic Average, implemented in the CLUSTERING CALCULATOR program developed by J. Brzustowski (<http://www.biology.ualberta.ca/jbrzusto/cluster.php>).

Null Alleles, Hardy-Weinberg Equilibrium

The existence and frequency of null alleles were tested (on the N dataset) using the Expectation Maximization (EM) algorithm [46] implemented in GENEPOP (http://genepop.curtin.edu.au/genepop_op8.html). Deviations from Hardy-Weinberg equilibrium (HWE) were tested using both alternative hypotheses of a deficit or of an excess of heterozygotes for each locus and across loci, at population and at cultivar levels using Fisher method implemented in GENEPOP (http://genepop.curtin.edu.au/genepop_op1.html). The Hardy-Weinberg Equilibrium tests were calculated based on both N and G datasets.

Genetic diversity

Rarefied allelic richness and private allelic richness were computed on the N dataset using the rarefaction method implemented in ADZE 1.0 [47], with a minimum sample size of $G=15$ genotypes. The small size of some samples when eliminating redundant MLGs from the populations limited the estimation of genetic diversity parameters. Genetic diversity was estimated, for each locus, at the population and cultivar levels. Values were estimated using

the unbiased expected heterozygosity $U_{He} = 2N / (2N-1) * He$ (where $He = 1 - \sum p_i^2$, and p_i is the frequency of the i th allele in the population) implemented in GENALEX, the observed heterozygosity H_o and the Wright's fixation index ($F_{is} = 1 - H_o / He$) [48], implemented in GENETIX 4.05.2 [49]. Differences in mean diversities for each parameter were compared between populations and between both cultivars using the Mann-Whitney U-test implemented in XLSTAT-Pro 7.5 (Addinsoft 1995-2015). The F_{is} over loci and populations were calculated using the N dataset to better describe the sampled data; we also calculated the F_{is} for populations with sample size (≥ 15 individuals) after elimination of the redundant genotypes (G datasets). Confidence intervals at 95% were calculated using 10 000 bootstraps (GENETIX software). Wilcoxon's signed rank test, implemented in XLSTAT-Pro 7.5, was performed for comparisons of heterozygosity expected under HWE and to test differences in F_{is} values based on N and G matrices between cultivars.

Population differentiation and organization of genetic diversity

The genetic differentiation among collected populations was calculated for the N dataset. Null alleles may lead to overestimate population genetic differentiations as measured by F_{st} [50]. To calculate pairwise genetic distances between each pair of populations we therefore used corrected F_{st} values using the ENA (excluding null alleles) method, and corrected values of the Cavalli-Sforza and Edwards [51] genetic distance (D_c) using the INA (including null alleles) method, implemented in FreeNA [50]. Geographic distances between populations were estimated based on the populations coordinates using the following formula: $\ln(1 + \text{Geographic Distance})$ implemented in GENALEX. In order to test isolation by distance (IBD) within each cultivar, Mantel tests [52] were conducted to measure Pearson's correlation using the GENALEX software; significance at $\alpha = 0.05$ was determined using 1000 randomizations of the elements of the genetic distance (D_c) x geographic distance matrix.

Two different approaches were used to infer the population structure of the whole dataset. The first was a Bayesian clustering approach assuming Hardy-Weinberg and linkage equilibrium, as implemented in the STRUCTURE 2.3.4 software [53]. This analysis benefited from the Montpellier Bioinformatics Biodiversity platform services (<http://mbb.univ-montp2.fr/MBB/subsection/onlineTools.php?section=2>). STRUCTURE was run on the whole data, using both the G and the N datasets, to infer population structure in K genetic clusters and assign individuals to those clusters. Structure was run also on partitioned subsets, for each cultivar, in order to investigate lower levels of structure. Individuals were assigned by probability to a cluster, or jointly to two or more clusters if their genotypes indicated that they were admixed. Individuals for which more than 85% of alleles were assigned to a cluster were considered to belong to a single cluster. To infer posterior probabilities of K, we ran a series of independent runs with different user-defined values of K ranging from 1 to 10. The analysis was run ten times for 10^7 iterations of MCMC after a burn-in period of 10^6 based on pre-defined population genetics admixture and no linkage models. To identify the number of K clusters explaining the observed genetic structure we used the statistic parameters defined by Evanno et al. [54] based on the rate of change in the log probability of data between successive K values, implemented in a web-based program, STRUCTURE HARVESTER 0.6.94 [55]. The CLUMPP [56] output served to DISTRICT 1.1 software [57] to graphically represent the estimated population structure, according to the geographic proximity and ecological region of each cultivar. Each individual was represented by a thick line, which was partitioned into K colored segments, representing the individual's estimated membership fractions in the K clusters.

The population genetics model in STRUCTURE, relies on restrictive explicit assumptions (populations at Hardy-Weinberg equilibrium and linkage equilibrium) which are likely to be violated for cultivated species, especially those exposed to clonal propagation. The presence

of substantial null allele frequencies may also affect the estimation of the population differentiation [50]. Hence, we choose the discriminant analysis (DA) of principal components (DAPC; [58]) as a second approach, as it is a non-model clustering method. DAPC was performed to identify and describe clusters of genetically related individuals from G dataset, using the Adegenet 2.0.1 package [59] implemented in R software [60]. This multivariate analysis seeks linear combinations of the original variables (alleles), which maximize between-group variation and minimize within-group variation [58].

DAPC is a multivariate method that relies on data transformation using a principal component analysis as a prior step to the DA. This prior transformation ensures that variables subjected to the DA are uncorrelated. The identification of genetic groups was achieved using K-means [61], a clustering algorithm which finds a given number (K) of groups maximizing the differentiation between groups. To identify the optimal number of groups, K-means was run sequentially with increasing values of K, in a loop of 20 repetitions, and the optimal number of K groups was identified by its lowest Bayesian Information Criterion (BIC) [58]. We executed the cross-validation to provide an objective optimization procedure for identifying the number of PC axes to retain, since the number of PC axes can have a substantial impact on the results of the global DAPC analysis [58]. The cross-validation was carried out using the function `xvalDapc` with 100 repetitions. The cross-validation method was also used to provide individuals' assignment to groups as well as a visual assessment of between-population differentiation and admixture. Based on the optimal number of clusters and the retained discriminant functions, the DAPC analysis was run and derived probabilities for each individual membership in each of the different groups. Admixed individuals were identified at a threshold of 85% of assignation to a group.

Results

Information related to SSR markers

The 15 SSR loci revealed a total number of 375 alleles, ranging from six alleles at the BPPCT036 locus to 41 at the CPDCT045 locus (see Additional file 2: Table S1). All SSR markers, except BPPCT007, were in Hardy-Weinberg disequilibrium ($F_{is} \neq 0$) for N and G datasets. At the whole dataset level, four out of 15 loci were supposed to exhibit null alleles; the estimated frequency of these alleles ranged from 15% for UDP98-409 up to 27% for CPDCT045. Seven loci showed null alleles with frequencies over 10% for ‘Khachabi’; among them, two loci (UDP97-401 and CPDCT045) exhibited also null alleles for ‘Halwani’ (see Additional file 2: Table S1).

In all, 105 pairwise comparisons between loci were conducted over all datasets without redundant MLGs (see Additional file 3: Table S2). The coefficient of correlation was $r = 0.75$ for the whole dataset. At the cultivar level, significant genotypic LD ($P < 0.05$) was identified for ‘Khachabi’ $r = 0.22$ (and 95% of pairwise comparisons were in LD) while for ‘Halwani’ $r = 0.43$ (and 74% of pairwise comparisons were in LD) (Tables 2; see Additional file 3: Table S2).

The 15 microsatellite loci allowed to distinguish 314 MLGs out of 331 individuals distributed over the 14 ‘Khachabi’ populations, whereas 195 distinct MLGs in a total of 284 sampled individuals were distributed over the 11 ‘Halwani’ populations (Table 2). Our results showed that the set of 15 SSR loci allowed an efficient estimation of the true number of MLGs; the total number of MLGs was reached for 14 loci for the whole dataset and already for 11 and 12 loci for ‘Khachabi’ and ‘Halwani’ respectively (see Additional file 4: Fig. S2).

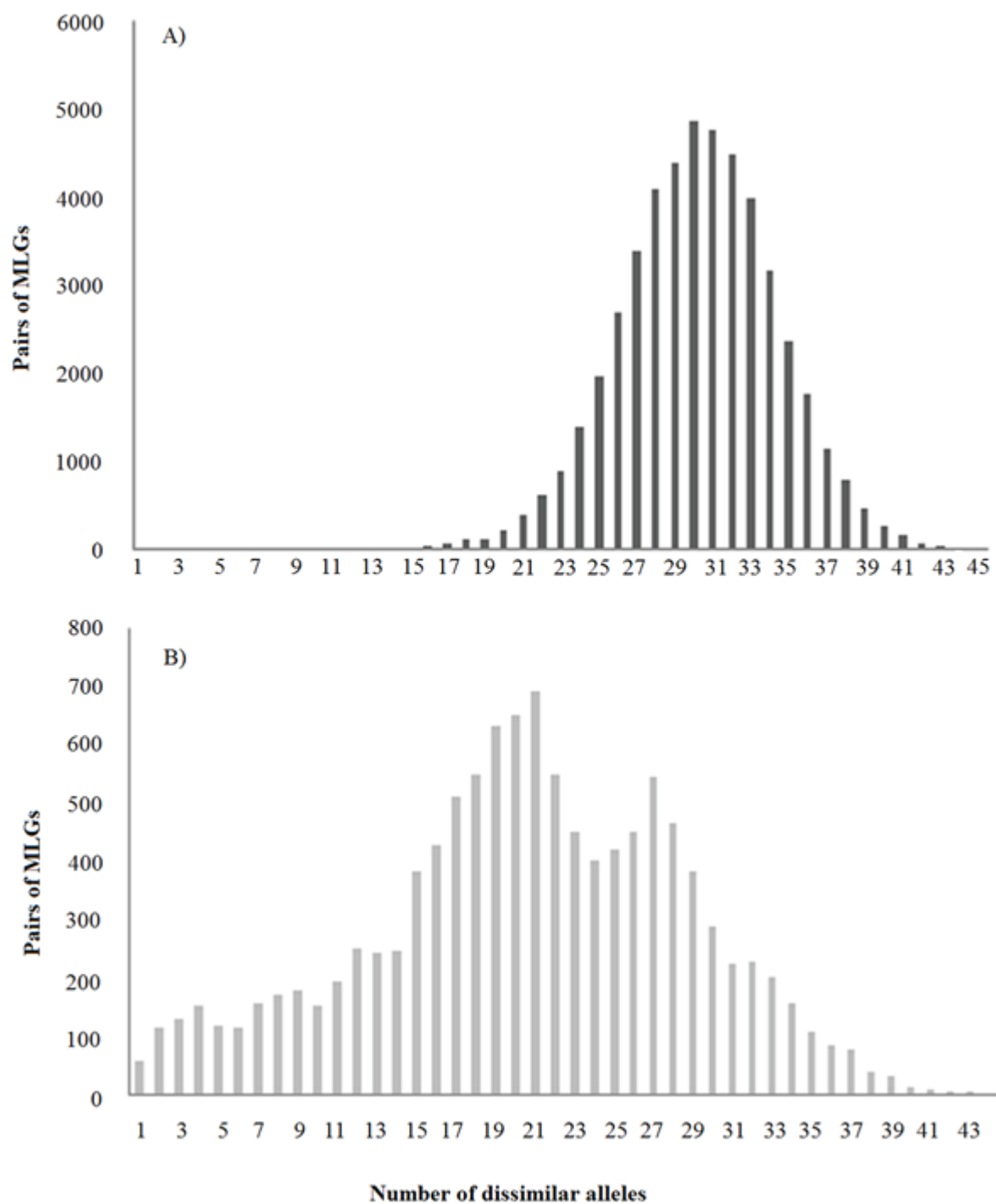


Figure 1 Distribution of the pairwise number of allele differences among MLGs.

A) Unimodal curve for 'Khachabi' genotypes; 48 826 pairwise comparisons, scale range from 0 to 6 000 B) A bimodal shaped curve for 'Halwani'; 11 264 pairwise comparisons after elimination of the 33 individuals assigned at 85% to 'Khachabi', Scale range from 0 to 800.

Genotypic diversity

The mean genotypic diversity value for cultivar Khachabi was significantly higher than the value reported for cultivar Halwani ($R_{\text{Khachabi}} = 0.95$ and $R_{\text{Halwani}} = 0.69$; U-test, $P = 0.000$; $\alpha = 0.05$). Populations from ‘Khachabi’ exhibited higher genotypic diversity levels than populations from ‘Halwani’, except for the ‘Khachabi’ KS_Ta population, due to the presence of three farmer-selected clones ($R = 0.59$) (Table 2). This indicates the presence of more redundant MLGs in ‘Halwani’ populations, as compared to ‘Khachabi’ populations. However, among ‘Halwani’ populations, HM_Am and HM_Je exhibited only two and one redundant MLGs respectively (corresponding $R = 0.93$ and 0.97 , respectively) and no redundant MLGs were identified in the HN_Bt population ($R = 1.00$).

We computed the distribution of the number of allelic differences between individuals (N dataset) for each cultivar; individuals were plotted after suppression of the redundant MLGs. In ‘Khachabi’, 48 798 pairwise comparisons ranged from 11 to 45 allelic differences with a peak at 30 differences, while only 28 pairwise comparisons were differing by 1 to 10 alleles (Fig. 1A). Individuals from cultivar Halwani were plotted after elimination of individuals assigned by STRUCTURE (see below) to the cluster that mainly comprised ‘Khachabi’ individuals (Fig. 1B). 11 264 pairwise comparisons for the ‘Halwani’ MLGs (Fig. 1B) displayed a continuity of values ranging from 1 to 43 differences with a multimodal curve; two major peaks, at 21 and 27, were observed. In comparison to ‘Khachabi’ a total of 1 542 pairwise comparisons differed by up to 10 alleles.

Forty-two redundant MLGs were defined as clones; three were observed in the clonally propagated ‘Khachabi’ population KS_Ta (data not shown) while 36 were observed in 10 of the 11 ‘Halwani’ populations. Among these, 24 were private (associated to one population)

Table 2 Genetic and genotypic diversity measurements.

| Pop | Genotypic Diversity | | | | | Genetic diversity | | | | | |
|----------|---------------------|-----|------|----------|---------------------|-------------------|-----------------|-------------|-------------|---------|-----------|
| | N | G | R | PI (N) | % LD _(G) | Ar ⁺ | Pa ⁺ | Ho (N) | He (N) | Fis (N) | Fis (G) # |
| Halwani | 284 | 195 | 0.69 | 1.20E-13 | 95 | 4.80 (0.36) | 2.01 (0.31) | 0.67 (0.07) | 0.66 (0.04) | -0.03 | 0 |
| HB_Bi | 23 | 8 | 0.32 | 4.40E-11 | 42 | 3.60 (0.25) | 0.02 (0.01) | 0.70 (0.08) | 0.62 (0.03) | -0.12 * | - |
| HB_Fe | 33 | 14 | 0.41 | 7.60E-08 | 77 | 2.89 (0.20) | 0.10 (0.07) | 0.62 (0.12) | 0.47 (0.06) | -0.34 * | - |
| HB_NA | 30 | 19 | 0.62 | 2.20E-12 | 83 | 3.93 (0.25) | 0.04 (0.02) | 0.74 (0.08) | 0.64 (0.05) | -0.15 | -0.08* |
| HB_Ze | 16 | 7 | 0.4 | 1.50E-06 | 14 | 2.12 (0.18) | 0.00 (0.00) | 0.54 (0.12) | 0.42 (0.06) | -0.28 * | - |
| HM_Am | 29 | 27 | 0.93 | 2.60E-08 | 20 | 2.86 (0.29) | 0.12 (0.08) | 0.63 (0.09) | 0.49 (0.06) | -0.33 * | -0.30* |
| HM_Ma | 30 | 21 | 0.69 | 6.10E-09 | 21 | 2.96 (0.24) | 0.05 (0.03) | 0.75 (0.09) | 0.52 (0.05) | -0.45 * | -0.39* |
| HM_Je | 30 | 29 | 0.97 | 1.60E-17 | 52 | 6.10 (0.40) | 0.21 (0.06) | 0.66 (0.04) | 0.76 (0.03) | 0.13 * | 0.14* |
| HN_Bc | 32 | 22 | 0.68 | 6.80E-11 | 70 | 3.65 (0.32) | 0.03 (0.01) | 0.75 (0.08) | 0.59 (0.05) | -0.28 | -0.21* |
| HN_Bt | 23 | 23 | 1 | 7.70E-12 | 43 | 3.91 (0.51) | 0.20 (0.09) | 0.62 (0.08) | 0.60 (0.07) | -0.03 | -0.03 |
| HS_DQ | 23 | 14 | 0.59 | 6.30E-08 | 50 | 2.93 (0.25) | 0.02 (0.02) | 0.73 (0.11) | 0.48 (0.06) | -0.56 * | - |
| HS_T a | 15 | 10 | 0.64 | 7.90E-09 | 18 | 3.10 (0.24) | 0.07 (0.03) | 0.52 (0.08) | 0.51 (0.06) | -0.06 * | - |
| Khachabi | 331 | 314 | 0.95 | 1.80E-25 | 74 | 8.28 (0.50) | 5.48 (0.45) | 0.68 (0.05) | 0.86 (0.03) | 0.20 | 0.20 * |
| KB_Ba | 26 | 26 | 1 | 2.30E-21 | 3 | 7.23 (0.47) | 0.23 (0.07) | 0.64 (0.05) | 0.84 (0.03) | 0.24 | 0.24 * |
| KB_Bi | 20 | 20 | 1 | 4.40E-16 | 21 | 6.08 (0.41) | 0.29 (0.08) | 0.70 (0.06) | 0.74 (0.03) | 0.05 | 0.05 * |
| KB_EM | 30 | 30 | 1 | 2.80E-21 | 13 | 7.47 (0.49) | 0.22 (0.06) | 0.73 (0.05) | 0.82 (0.03) | 0.12 | 0.12 * |
| KB_Ir | 15 | 15 | 1 | 3.70E-18 | 33 | 6.31 (0.47) | 0.41 (0.14) | 0.72 (0.06) | 0.79 (0.04) | 0.10 * | 0.10 * |
| KB_KL | 30 | 29 | 0.97 | 1.50E-20 | 13 | 7.01 (0.46) | 0.26 (0.07) | 0.70 (0.05) | 0.83 (0.02) | 0.15 | 0.15 * |
| KB_Lu | 30 | 30 | 1 | 1.10E-20 | 14 | 6.92 (0.43) | 0.31 (0.06) | 0.73 (0.05) | 0.83 (0.03) | 0.13 | 0.13 * |
| KB_Sh | 18 | 18 | 1 | 1.40E-17 | 37 | 6.19 (0.37) | 0.23 (0.11) | 0.63 (0.06) | 0.78 (0.04) | 0.18 | 0.18 * |
| KN_As | 16 | 16 | 1 | 3.10E-19 | 20 | 6.92 (0.53) | 0.41 (0.12) | 0.62 (0.07) | 0.80 (0.04) | 0.22 | 0.22 * |
| KN_EQ | 16 | 16 | 1 | 8.20E-19 | 14 | 6.83 (0.52) | 0.56 (0.13) | 0.62 (0.06) | 0.79 (0.05) | 0.22 * | 0.22 * |
| KS_Ba | 30 | 30 | 1 | 7.20E-21 | 22 | 7.20 (0.48) | 0.32 (0.08) | 0.71 (0.06) | 0.82 (0.03) | 0.12 | 0.12 * |
| KS_Bl | 29 | 29 | 1 | 2.40E-19 | 18 | 6.71 (0.42) | 0.18 (0.06) | 0.67 (0.05) | 0.80 (0.03) | 0.16 | 0.16 * |
| KS_Ch | 23 | 21 | 0.91 | 1.30E-20 | 25 | 7.14 (0.47) | 0.37 (0.12) | 0.71 (0.05) | 0.83 (0.03) | 0.14 | 0.15 * |
| KS_MJ | 18 | 17 | 0.94 | 2.90E-17 | 19 | 5.91 (0.40) | 0.12 (0.05) | 0.71 (0.06) | 0.77 (0.04) | 0.08 | 0.07 * |
| KS_T a | 30 | 18 | 0.59 | 1.80E-15 | 52 | 5.32 (0.30) | 0.13 (0.04) | 0.65 (0.08) | 0.73 (0.03) | 0.11 * | 0.15* |

N, Number of samples; G, Number of distinct genotypes; R, Proportion of distinct genotypes = (G-1)/(N-1); PI, probability of identity of two genotypes; %LD: percentage of genotypic linkage disequilibrium based on a single MLG copy; Ar⁺, average number of alleles per locus; Pa⁺, mean number of private alleles per locus; Ho, observed heterozygosity; He, unbiased expected heterozygosity; Fis, inbreeding coefficient.

⁺ Standardized population at G = 15 using rarefaction method.

Fis calculated for sample size over 15 individuals.

* Significant departure from Hardy-Weinberg equilibrium ($P \geq 0.05$).

X_(N) Values calculated based on N matrix

X_(G) Values calculated based on G matrix.

Standard errors are in brackets.

and 12 clones were found across different populations. Some clones were shared among geographically distant populations of ‘Halwani’.

Genetic diversity

To better describe the genetic diversity existing within each cultivar and population, we decided to conduct analyses using the N dataset. ‘Khachabi’ exhibited a larger number of alleles than ‘Halwani’ with 361 and 221 alleles, respectively (Table 2).

Genetic diversity parameters were calculated over loci and the values are presented with their standard errors (Table 2). The mean rarefied allelic richness (Ar) and the private allelic richness (Pa) were significantly higher (U-test, $P < 0.0001$; $\alpha = 0.05$) for ‘Khachabi’ (Ar = 8.28 ± 0.50 and Pa = 5.48 ± 0.45), as compared to ‘Halwani’ (Ar = 4.80 ± 0.36 and Pa = 2.01 ± 0.31). The mean genetic diversity was significantly higher for ‘Khachabi’ (He = 0.86 ± 0.03) than for ‘Halwani’ (He = 0.66 ± 0.04) (U-test, $P = 0.002$; $\alpha = 0.05$). Cultivar Khachabi also displayed a significant deficit in heterozygosity (Fis = 0.20) while cultivar Halwani had no deficit in heterozygosity (Fis = -0.03). The comparison of the fixation index (Fis) (N dataset) exhibited high differences between ‘Khachabi’ and ‘Halwani’ (Wilcoxon’s test, $P < 0.0001$; $\alpha = 0.05$). The confidence intervals of Fis values are presented in Table S3 (see Additional file 5).

Different genetic diversity levels were also observed among populations, the lowest number of alleles for the ‘Halwani’ populations being in HB_Ze (Ar = 2.12 ± 0.18) and the highest number in the HM_Je population (Ar = 6.10 ± 0.40) (Table 2). The number of private alleles (Pa) ranged from 0 to 0.21. Ten out of eleven ‘Halwani’ populations exhibited an excess of heterozygotes with a significant departure from Hardy-Weinberg equilibrium ($P \geq 0.05$). The number of rarefied private alleles (Pa) ranged from 0.12 to 0.56. All ‘Khachabi’ populations exhibited a high deficit in heterozygosity ($P \leq 0.05$) (Table 2).

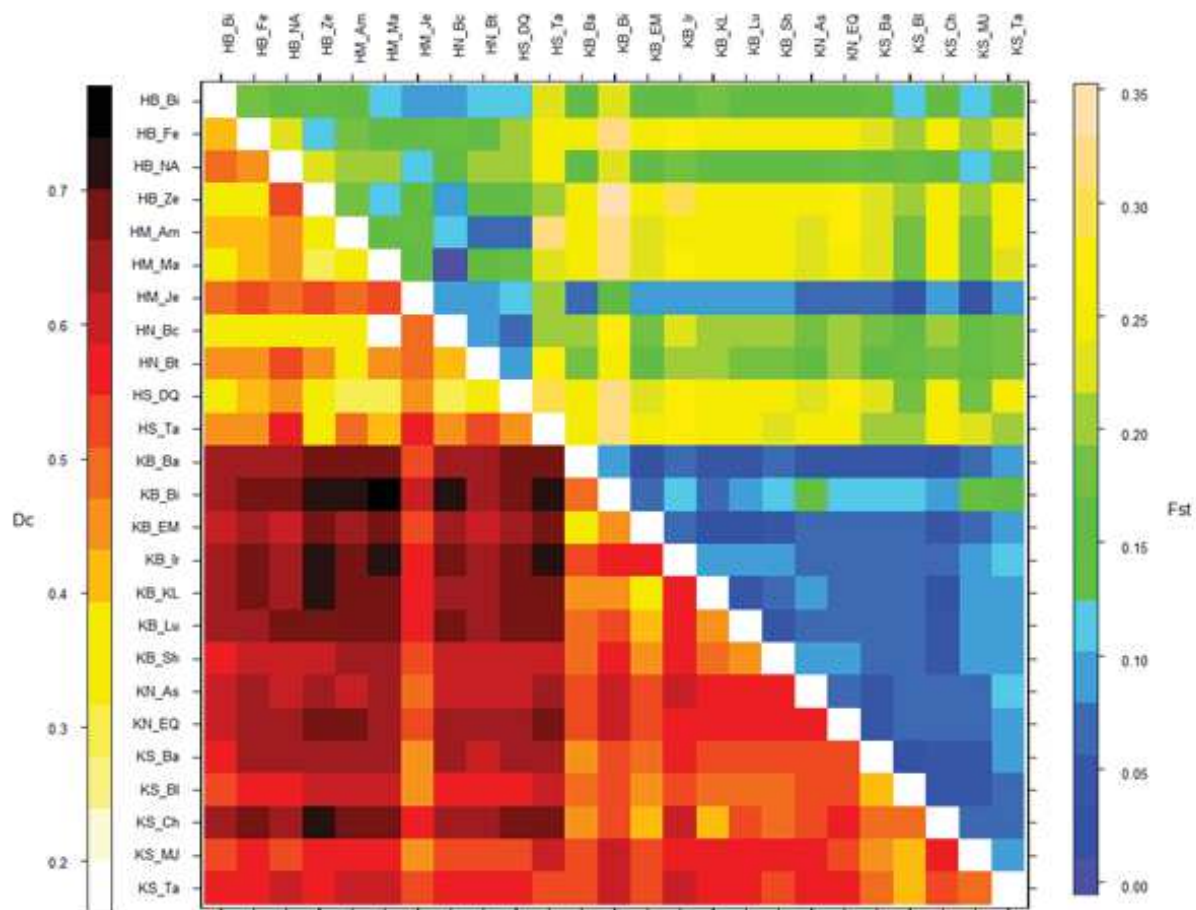


Figure 2 Heat map for pairwise genetic distance between populations.

Upper part: Fst; lower part: Dc (Cavalli-Sforza). Values were calculated with 1 000 permutations. Values are significant at $P > 0.001$. Genetic differentiation is: Low to moderate $0.02 < x < 0.15$ (in blue), high genetic differentiation $0.15 < x < 0.25$ (in green) and very high > 0.25 (in yellow). Genetic distance (Dc) is: Low to moderate $0.02 < x < 0.4$ (white to yellow), high genetic distance $0.4 < x < 0.8$ (red to black).

Population genetic structure

Genetic differentiation (F_{st}) varied from 0.03 to 0.12 among ‘Khachabi’ populations, and from 0.01 to 0.31 among ‘Halwani’ populations; the highest F_{st} observed among populations from both cultivars was 0.33 (between HB_Ze and KB_Bi) (Fig.2). The pairwise estimate values of D_c varied from 0.37 to 0.62 among ‘Khachabi’ populations, and from 0.20 to 0.56 among ‘Halwani’ populations, and from 0.44 to 0.74 among populations from both cultivars (Fig.2). The Mantel test revealed a pattern of isolation-by-distance in ‘Khachabi’ ($R^2 = 0.20$, $P = 0.006$) but not in ‘Halwani’ ($R^2 = 0.04$, $P = 0.124$) (see Additional file 6: Fig. S3).

The STRUCTURE analysis on the overall G dataset revealed the highest ΔK for $K = 2$ ($\Delta K = 1640.26$, $H' = 0.999$); 92% of ‘Khachabi’ individuals were assigned to cluster 1 (in red) and 81% of ‘Halwani’ individuals were assigned to cluster 2 (in green) (Fig. 3A). Admixed individuals among both clusters were observed in both cultivars at a threshold of assignment of 85%. The lowest percentages of assignment to the cultivar Halwani were observed in HB_Bi, HB_NA and HM_Je populations with 57%, 58% and 34% of individuals assigned to cluster 2, respectively. On the whole sample, the STRUCTURE analysis assigned 117 ‘Halwani’ individuals, belonging to 29 clones, to cluster 2. Fourteen individuals from HB_NA belonging to five clones and two individuals from the HB_Fe ‘Halwani’ populations, belonging to one clone, were assigned to cluster 1. Eight individuals from HB_Bi belonging to two clones and three individuals from the HB_NA belonging to one clone were admixed among both clusters.

The STRUCTURE analysis of ‘Khachabi’ assigned its 314 MLGs into two clusters ($K = 2$, $\Delta K = 7024.5$ and $H' = 0.999$) (Fig. 3B). Populations belonging to these clusters were

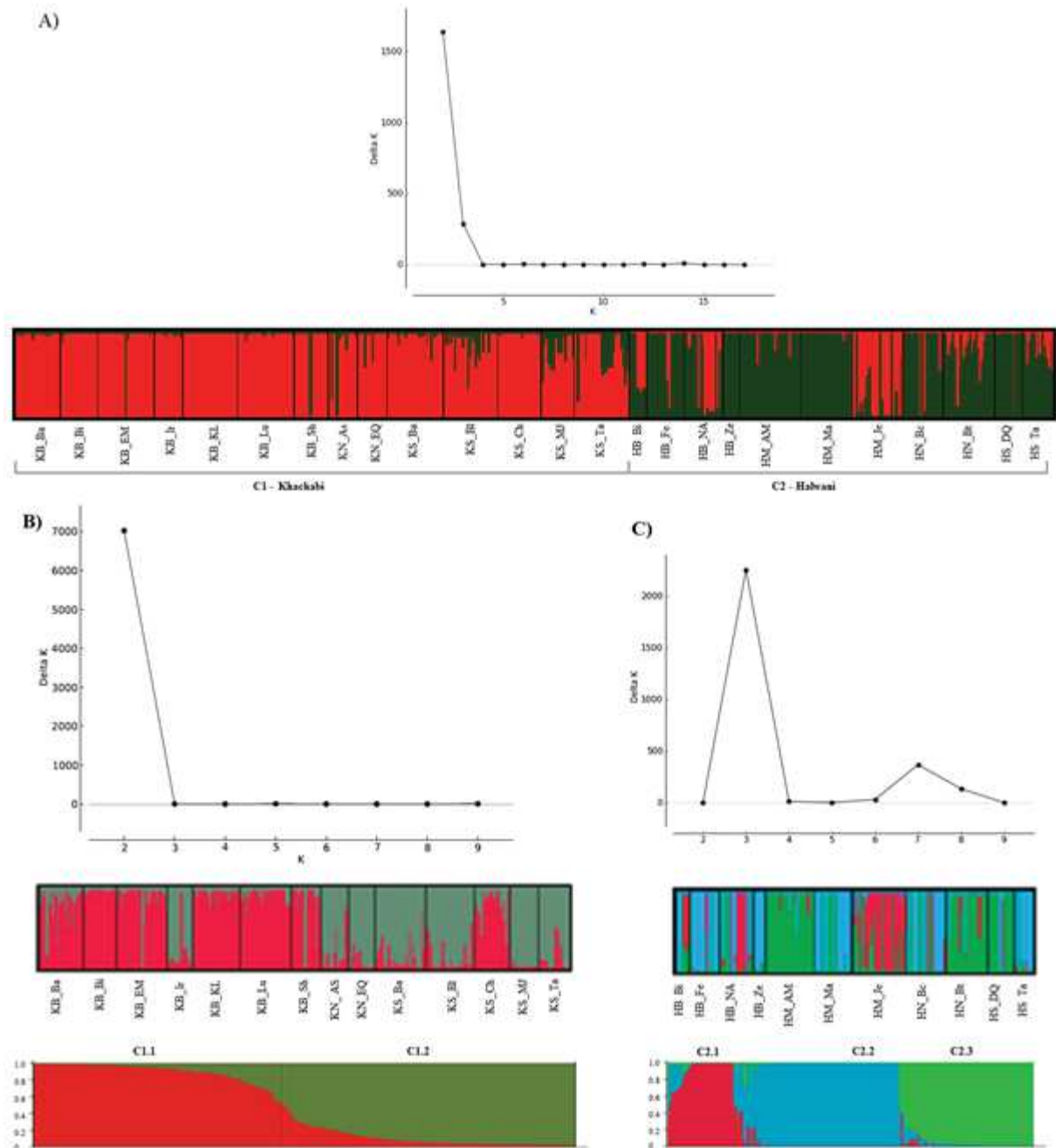


Figure 3 Genetic Structure of almond genotypes estimated by STRUCTURE.

A) DISTRUCT representation of 14 'Khachabi' and 11 'Halwani' populations (509 genotypes). $K = 2$, $\Delta K = 1640.26$ and $H' = 0.998$; B) DISTRUCT representation of 14 'Khachabi' populations (314 genotypes); ΔK graph: $K = 2$, $\Delta K = 7024.5$ and $H' = 0.999$; C) DISTRUCT representation of 11 'Halwani' populations (195 genotypes); $K = 3$, $\Delta K = 2251.1$ and $H' = 0.999$.

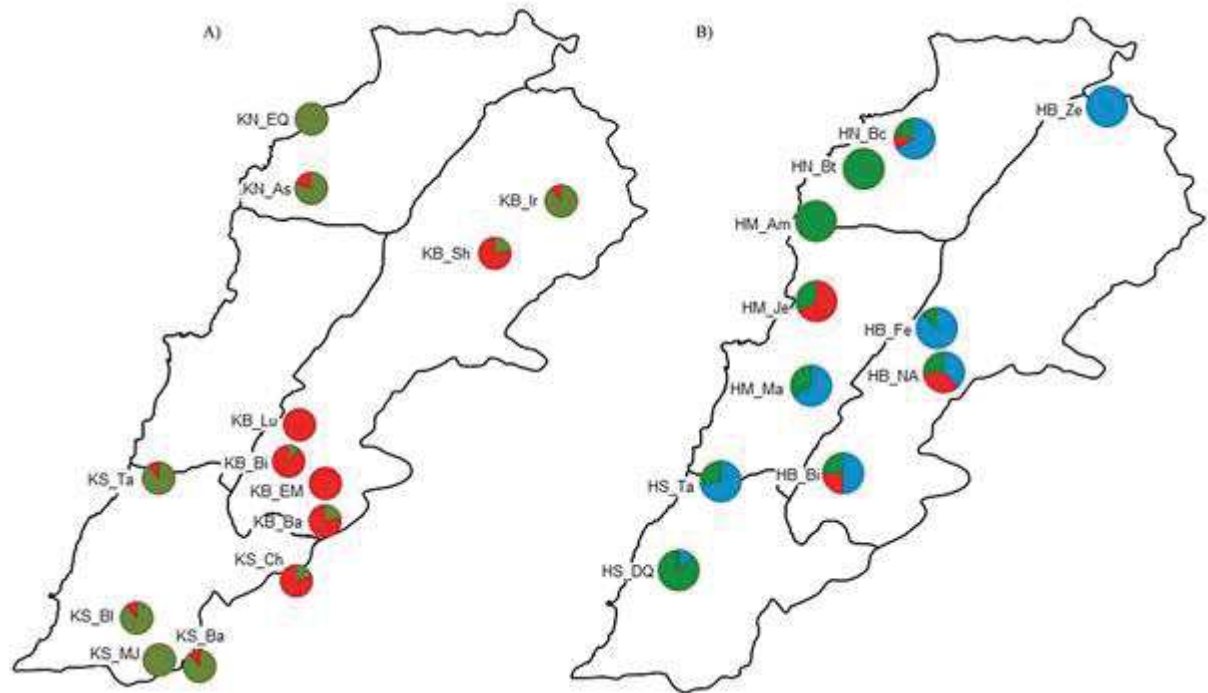


Figure 4 Geographical distribution of STRUCTURE assignment, each population is represented in a pie chart.

A) $K = 2$ for 'Khachabi' populations; B) $K = 3$ for 'Halwani' populations.

Full names of the abbreviations for the populations are shown in Table 1.

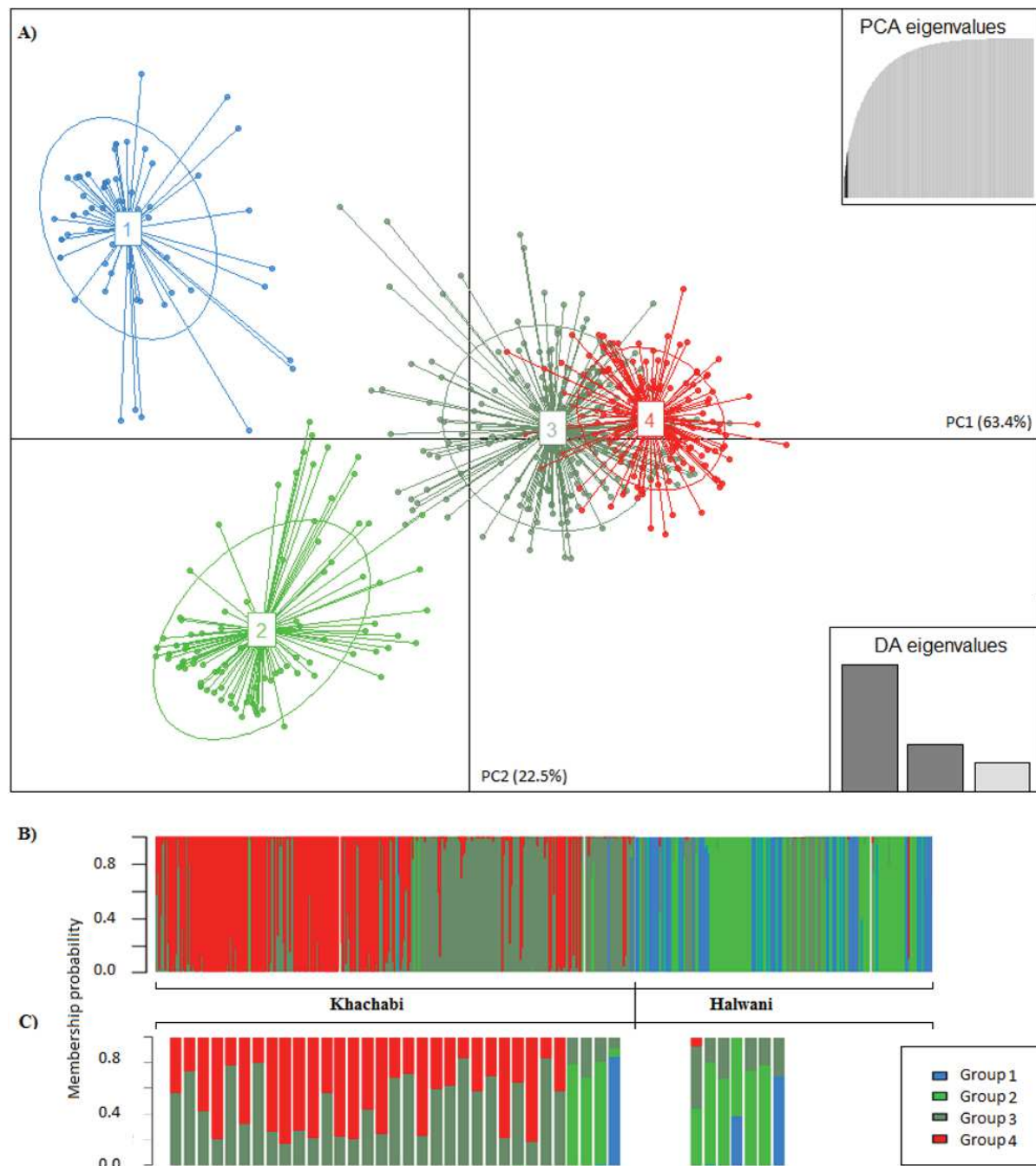


Figure 5 Genetic structure of almond genotypes estimated by DAPC.

A) Scatterplot of 509 MLGs on the first two axes representing 63.4% (PC1) and 22.5% (PC2) of the variation. Individuals are represented as dots and the groups as inertia ellipses. B) STRUCTURE-like plot of 509 MLGs; C) Admixed individuals having no more than 85% of probability of membership in a single group. Each line corresponds to a single individual; colors correspond to the group memberships.

geographically differentiated (Fig. 4A). The first cluster (C1.1, red) comprised the populations of the Bekaa region, except for the population KB_Ir; the nearby population from the South, KS_Ch, belonged also to this cluster. The second cluster (C1.2, olive green) was composed of populations from the South and North regions and the KB_Ir population from the Bekaa region. Some individuals were admixed among both clusters with percentages of assignment under 85%.

The STRUCTURE analysis of the populations named as ‘Halwani’ assigned the 195 MLGs of this cultivar into three clusters ($K = 3$, $\Delta K = 2251.1$ and $H' = 0.999$) (Fig. 3C). As expected, the first cluster (C2.1, red) included the individuals which were re-assigned to ‘Khachabi’ in the global analysis (C1, Fig. 3A). The other individuals were assigned into two clusters (C2.2 in blue and C2.3 in green, Fig. 3C). The geographic differentiation of populations was not as clear-cut as for ‘Khachabi’ populations (Fig. 4B). Individuals from a same population were mainly assigned to the same cluster. Individuals from HN_Bt and HM_AM exhibited admixture between both clusters.

The likelihood-based analysis of population structure using DAPC split the 509 MLGs from the 25 populations into four distinct groups (Fig. 5A and B). The ‘Khachabi’ MLGs were assigned to P3 (150 MLGs) and to P4 (153 MLGs), except for four MLGs assigned to the P1 and seven to P2 groups (Table 3). At a threshold value of 85% of assignment, a small number of individuals showed a pattern of admixture: 10 among P3, 19 among P4, 4 among P1 and 5 among P2 (Fig. 5C). The ‘Halwani’ MLGs were assigned into two more differentiated DAPC groups (Table 3): P1 (63 MLGs) comprised mainly individuals from the HM_Ma, HN_Bc and HB_Fe populations; P2 included 94 MLGs from all populations except HB_Ze, 37 MLGs were assigned to P3 and one MLG to P4. With the threshold value of 85% of assignment, only two MLGs of the P1 and three of the P2 were admixed. Admixed individuals were identified within and among cultivars (Fig 5C). Within cultivars, 29 ‘Khachabi’ individuals were

Table 3 Assignment of 509 sampled individuals to STRUCTURE's clusters and DAPC's groups.

| STRUCTURE | C1 | | C2 | | Admix* between C1 and C2 | | | | G |
|-----------|--------|--------|--------|----|--------------------------|-------|-------|-------|-----|
| DAPC | P3 | P4 | P1 | P2 | P1 | P2 | P3 | P4 | |
| HB_Bi | | | 3 | 1 | | | 3 | | 7 |
| HB_Fe | 1 | | 12 | 2 | | 1(1)* | | | 16 |
| HB_NA | 7 | | 4 | 4 | | 4(1)* | | | 19 |
| HB_Ze | | | 7 | | | | | | 7 |
| HM_Am | | | | 24 | | 3 | | | 27 |
| HM_Ma | | | 13(1)* | 8 | | | | | 21 |
| HM_Je | 14 | | | 4 | | 6 | 5 | | 29 |
| HN_Bc | 2 | | 13 | 6 | | | 1 | | 22 |
| HN_Bt | | | 1(1)* | 11 | 1 | 8(1)* | 2 | | 23 |
| HS_DQ | | 1 | 2 | 11 | | | | | 14 |
| HS-Ta | | | 5 | 1 | 2 | | 2(1)* | | 10 |
| KB_Ba | 8(3)* | 18(4)* | | | | | | | 26 |
| KB_Bi | | 20 | | | | | | | 20 |
| KB_EM | 2 | 27(2)* | | 1 | | | | | 30 |
| KB_Ir | 11 | 4(1)* | | | | | | | 15 |
| KB_KL | 1(1)* | 27(1)* | | | | | | | 28 |
| KB_Lu | 6(1)* | 24(3)* | | | | | | | 30 |
| KB_Sh | 3(1)* | 13(3)* | 1 | 1 | | | | | 18 |
| KN_As | 13(1)* | | | 1 | | 1 | 1 | | 16 |
| KN_EQ | 13(1)* | 1 | | | | | 2 | | 16 |
| KS_Ba | 28(1)* | 1 | | | | | 1 | | 30 |
| KS_Bl | 18(1)* | 2(2)* | | | | | 8 | 1 | 29 |
| KS_Ch | 7 | 14(1)* | | | | | | | 21 |
| KS_MJ | 6 | | | | | 3(2)* | 8 | | 17 |
| KS-Ta | 5 | | 1(1)* | | 2(1)* | | 9 | 1(1)* | 18 |
| Total | 145 | 152 | 62 | 75 | 5 | 26 | 42 | 2 | 509 |

* <85% of alleles were assigned to a cluster and/or group.

x (y), x: number of assigned individuals; y: number of admixture within assigned individuals, resulting from DAPC analysis.

G, Number of distinct genotypes.

admixed between P3 and P4 groups while one ‘Halwani’ individual was admixed between P1 and P2 groups. Among cultivars, one ‘Khachabi’ individual was admixed among P1, P2 and P3, three MLGs were admixed between P2 and P3, one ‘Halwani’ MLG from HS_Ta was admixed between P1 and P3, another MLG was admixed among P2, P3 and P4, and four other ‘Halwani’ MLGs were admixed between P2 and P3.

The results from both Bayesian and non-model based analyses were compared at a threshold of 85% of assignment, for the global dataset (both cultivars; Table 3). Among ‘Khachabi’ individuals, 78% were assigned to cluster C1 according to STRUCTURE and to the P3 and P4 DAPC groups. As for ‘Halwani’, 66.7% of individuals assigned to cluster C2 according to STRUCTURE were assigned to the P1 and P2 DAPC groups.

Considering cultivars separately, the ‘Khachabi’ C1.1 and C1.2 STRUCTURE clusters correspond to P3 and P4 DAPC groups, respectively. As to ‘Halwani’, individuals assigned to C2.1 by STRUCTURE were assigned to P3 by DAPC; individuals assigned to the C2.2 and C2.3 STRUCTURE clusters were mainly assigned to the P1 and P2 DAPC groups, respectively.

Discussion

Understanding how human practices affect the evolutionary dynamics of domesticated plant species is a key factor to improve the management of crop genetic diversity. It also helps understand how specific biological parameters, like reproductive systems and gene flow, affect genetic diversity. The cultivated almond is intermediate between seed propagated and vegetatively propagated perennial species; it provides a good model to answer our inquiries on domestication and diversification of vegetatively reproduced plants cultivated for their fruits and particularly self-incompatible fruit trees. Indeed, almond has been domesticated through seed selection and reproduction [9]; grafting was developed only over the last

centuries; and its self-incompatibility system interferes with its sexual reproduction, as well as with fruit production *per se*. We will discuss our results concerning: 1) the high genetic diversity level of ‘Khachabi’ and its lower level in ‘Halawani’; 2) the deficit of heterozygosity in ‘Khachabi’ and its excess in ‘Halwani’; 3) the high genetic differentiation among ‘Halwani’ and ‘Khachabi’; 4) the high genetic differentiation between ‘Halwani’ groups; 5) the low genetic differentiation between both ‘Khachabi’ groups.

Genetic diversity

As expected from the propagation mode of each cultivar and the outcrossing mating system of the species, the sexually propagated cultivar Khachabi exhibited much more genetic and genotypic diversity than the vegetatively propagated cultivar Halwani.

The level of genetic diversity observed within ‘Khachabi’ ($H_e = 0.86$) was as high as what was observed in national germplasm collections of different Mediterranean countries [19, 20 and 22]. This genetic diversity could even be compared to the genetic diversity of spontaneous *Prunus* species, such as *P. orientalis* sampled from Lebanon, Syria and Turkey ($H_e = 0.85$) [26]. The comparison holds also at the population level. Indeed, in all seed-propagated ‘Khachabi’ populations, genetic diversity ranged between 0.74 and 0.84. These values compare with those observed in 20 wild Chinese populations of *P. sibirica* L., another self-incompatible species, where H_e ranged from 0.66 to 0.80 [27]. In our sample, the genotypic diversity level was only reduced for the ‘Khachabi’ KS-Ta population ($R = 0.59$), where the farmer had selected, and graft-propagated, three trees selected for their fruits. Even so, this unusual situation did not affect the level of allelic richness, as the selected individuals had retained a level of heterozygosity as high as for the seed-propagated ‘Khachabi’ populations.

‘Halwani’ exhibited a lower genetic diversity than ‘Khachabi’; which is due to the introduction of a limited number of individuals followed by clonal propagation. However, genetic diversity of ‘Halwani’ populations is still relatively high (H_e from 0.42 to 0.59), compared to cultivated self-pollinated *Prunus* species, as in the examples of three genetic groups of commercial peach cultivars collected from North America and Europe (H_e from 0.36 to 0.43) [38] and two genetic groups of sweet cherry cultivars collected from 16 countries and maintained in France (H_e values of 0.27 and 0.30, respectively) [62]. We relate this relatively high diversity to self-incompatibility in cultivar Halwani. Indeed, if orchards were constituted by a single self-incompatible genotype, flower fertilization would be hampered and commercial fruit production reduced. The only solutions are either to maintain pollen flow from external sources, or to maintain a minimal diversity of self-incompatibility alleles within ‘Halwani’ orchard. The first solution is observed when ‘Khachabi’ or spontaneous almond trees are present in the close vicinity, so they can participate in pollination. The second one implies a conscious choice of genetically diverse grafting materials, to avoid mate limitations within ‘Halwani’ orchards. Indeed, all ‘Halwani’ populations presented a composite structure with different genotypes. These management practices are found in cultivars of other *Prunus* species. For example, similar traditional management has long been known in cultivated apricot, where much of the variability has been conserved using a mix between seed- and graft-propagation (e.g. Vesuvian, Roussillon, and Peloponnese apricots) [63].

‘Khachabi’ exhibited a high heterozygosity deficit at both cultivar and population levels. These deficits can be explained in part by the presence of null alleles in seven loci, whose frequencies were significantly different from zero. Null alleles lead to underestimate the observed heterozygosity within populations and overestimate the fixation index [64]. It is

likely that, on the long term, some level of endogamy (continued crosses among trees within population) has also contributed to this deficit.

The heterozygote excesses observed in ‘Halwani’ populations were in accordance with expectations; as repeated clonal propagation has preserved the initial heterozygosity of founder individuals. Similar results were observed for the perennial outcrossing crop cassava in Brazil [65], for improved apricot in Spain [66] and for *Prunus avium* in three French sites [8]. Over generations, somatic mutations may also contribute to increase heterozygosity. Such mutations were evidenced in pairwise comparisons among collected individuals: some MLGs were differentiated from others at only one of the fifteen analyzed loci, and by slight inter-allelic variations (1 or 2 bp).

Genetic differentiation

The exploration of genetic structure by using both Bayesian and non-model methods has provided complementary information, by identifying different levels of genetic structure within the overall dataset (Fig. 5, Tables 3). The first, global, STRUCTURE analysis differentiated among cultivars while DAPC analysis detected structure among and within cultivars. The latter sub-structuring was also obtained when carrying out the STRUCTURE analysis for each cultivar separately. It is worth noting that the value of the statistic ΔK for $K = 2$ was large enough to approve the robustness of these clusters, and exclude the effect of some departure from STRUCTURE model hypotheses that may lead to a difficult estimation of the true number of clusters [67]. The assignment of individuals to DAPC groups was consistent with their assignment to the Structure clusters; we observed only a slightly lower consistency in the identification of admixed individuals between both methods.

Clustering results were in good accordance with cultivar denominations; the two major gene pools (STRUCTURE clusters C1 and C2) correspond to the studied cultivars, and the

assignment of individual trees to one of them mostly reflects their denomination by farmers. Nevertheless, history of orchard establishment and rejuvenation has produced some discordance between farmers' denominations and the genetic identity of some trees: 30 'Halwani'-named trees, from three different orchards, are assigned to the 'Khachabi' (C1) gene pool. This could be due either to the presence of 'Khachabi' trees in the 'Halwani' planting material coming from the nursery or to a failure of 'Halwani' grafting onto 'Khachabi' old trees when rejuvenating the orchard.

Dominant clonal propagation of cultivar Halwani has led to limited gene flow with 'Khachabi'. Only four 'Khachabi' and six 'Halwani' individuals were not assigned to any groups by DAPC, and the detection of admixture suggests that these 10 genotypes were produced by hybridization. This low level of admixture was the result of a combination of both sexual and clonal propagation practices: sowing spontaneous hybrids between both cultivars, followed by clonal propagation of selected hybrid trees; this ensured bi-directional gene flow between cultivars. This effect of the propagation mode on the genetic differentiation was also documented for cultivated apricot: even though apricot in Tunisia has experienced a single introduction event from the Irano-Caucasian area, it has been subject to local diversification as a result of different propagation practices that have led to the existence of two different gene pools, one propagated by seeds and the other one propagated by grafting [21].

Our results demonstrate also that 'Khachabi' is structured into two relatively close gene pools ($F_{st} = 0.033$) separated geographically by the chain of Lebanese mountains; they correspond mainly to populations of the central plateau of the Bekaa region and its back mountains, for the first gene pool, and populations of the maritime range of coast and adjacent mountains, for the second gene pool. Admixed individuals were observed between these two gene pools. Farmers mainly established their 'Khachabi' orchards with seeds from nearby orchards or

mountains (presumably ancient areas of cultivation). However, the presence of individuals from both gene pools, as well as admixed individuals, in the same orchards, attests for events of long distance seed migration among different geographic regions (Fig. 4A). Recurrent hybridizations between ‘Khachabi’ gene pools as shown by the sigmoidal shape of proportion of assignation among both ‘Khachabi’ clusters have weakened their differentiation. Likewise, seed migrations and hybridizations among individuals from nearby populations have contributed to reduce the effects of isolation by distance among collected populations. The first ‘Khachabi’ gene pool (C1.1) had no admixed individuals with either of the ‘Halwani’ gene pools. The clonal propagation of ‘Halwani’ and its recent introduction in the Bekaa region have prevented hybridizations. The second ‘Khachabi’ gene pool shares the same geographic range as the ‘Halwani’ cultivar, and longtime coexistence and sexual reproduction have resulted in admixture between cultivars.

As to cultivar Halwani, our results indicate that two different gene pools ($F_{st} = 0.152$) were introduced. This double origin was supported by the high values of F_{st} and D_c observed between some ‘Halwani’ populations. The first gene pool (C2.3) is mainly present in old orchards (Table 1). These orchards included more admixed individuals, confirming that propagation after these early introductions has involved sexual recombination. The second gene pool (C2.2) is mainly present in the newest orchards (established in the 1980’s and 1990’s according to farmer interviews) where almond grafting was more intensively used. Therefore, it is not surprising that we observe only one admixture event between both gene pools, as time has been too short for consequent hybridization, particularly as concomitant evolutions of farmer’s practices have more clearly favored vegetative propagation over sexual recombination in the C2.2 gene pool of this cultivar. However, some pairs of geographically distant orchards, assigned to the same gene pool, exhibited low genetic distances, suggesting

that a network of exchanges between farmers was effective on a large scale and/or that farmers have brought clonal plant material from a common source.

Evolution of cultivated almond

The genetic diversity of the almond cultivars collected from this center of diversification revealed patterns of evolution, genetic structure, and geographical and social factors that have affected the crop in its evolutionary dynamics.

Evolutionary dynamics of 'Khachabi'

In 'Khachabi', genetic structure patterns were similar to those observed for wild tree species. However, gene flow is not random among 'Khachabi' populations. Almond pollination is mainly carried out by bees. Although, in case of winter food scarcity, they can carry pollen between trees at distances up to 5 km [68], such long distance pollen exchanges are exceptional since this cultivar and other cultivated almond trees are present everywhere. Instead, local dispersal of seeds (by farmers or animals) and pollen explain the limited differentiation among neighboring populations and the correlation between genetic and geographic distances that we have observed (see Additional file 6: Fig. S3); this is particularly true among the populations of the western Bekaa and the nearby Chebaa in the South (KS_Ch, Fig. 2 and 4). A limited isolation-by-distance is also consistent with a low level of endogamy and the consequent deficit in heterozygosity.

Dynamics of the integration of 'Halwani' grafted almond, and the adoption of grafting practices

'Halwani' was introduced from two different gene pools and propagated mainly in the coastal region. Both gene pools were introduced into the coastal regions by single or multiple introductions. The earliest one was concomitant with the incorporation of grafting practices

for almond. Based on case studies from our data, we can envisage a consistent scenario explaining the evolutionary dynamics of ‘Halwani’, and particularly its interactions with ‘Khachabi’. When ‘Halwani’ was introduced in a given orchard, individuals of ‘Khachabi’ were preserved, leading to hybridization between both cultivars and the advent of admixed MLGs. Indeed, non-admixed ‘Halwani’ material sampled in the oldest ‘Halwani’ orchard (HM_Je) was only represented by three individual trees. As the subsequent phase of diffusion of ‘Halwani’ was based on locally collected materials, admixed materials could be part of the process. Thus, in the HB_Bi orchard, even though the farmer was sure to have ‘Halwani’ trees (having bought grafted trees at the local nursery), our genetic analysis identified trees with admixed MLGs resulting from hybridization between both cultivars. The presence of clones with admixed MLGs in the same geographic region further supports a scenario of propagation from nurseries. Another likely source of error in the diffusion of ‘Halwani’ was grafting failure, followed by ‘Kachabi’ rootstock suckering. This may be suspected for the HB_Na orchard; the farmer had bought his plants from a local nursery as ‘Halwani’, but the genetic analysis assigned many of them to ‘Khachabi’.

Later on, the grafting technique was better adopted. The success of grafting increased, explaining the marked predominance of clones in the C 2.2 gene pool that was more recently introduced to the Bekaa region (HB_Ze and HB_Fe orchards). The reduction of on-farm sexual propagation and the corollary appropriation of almond grafting practices led to a reduction of the genetic diversity of ‘Halwani’. Furthermore, farmers do not propagate all clones at the same rate, and in the long term, some clones will be lost.

In addition to the effect of farmers’ practices, networks of exchange have played an important role in structuring genetic diversity. Indeed, gene flow appeared to result from a network of exchanges of grafting material when establishing new orchards. The ‘Halwani’ grafts may have been taken from nurseries and/or orchards. This history of introduction and the exchange

network of both 'Halwani' gene pools deserves to be explored more in depth in order to better understand the impact of human practices on the dynamics of this cultivar and its interactions with more ancient local germplasm.

Evolution of domesticated almond and propagation practices

Our comparative study of the two main Lebanese almond cultivars and their interactions provided us with the main elements of a model of evolution of self-incompatible perennial fruit crops under domestication and traditional cultivation. Furthermore, the comparison allowed us to describe the genetic effects of the concomitant introduction of a new cultivar and the corresponding propagation techniques. Grafting techniques are recent in cultivated almond evolution, and the first phases of its domestication process only involved allogamous sexual propagation; the late adoption of vegetative propagation limits the interest of the parallels often established with the domestication of crops derived from wild relatives that propagate vegetatively, such as cactus pear, banana and pineapple.

As long as allogamous sexual propagation is exclusive, genotypes cannot be fixed, and the genetic structure of cultivated populations remain similar to that of the fruit crop wild relatives, showing high genetic diversity within populations, limited differentiation among populations, and isolation by distance and topography. Thus, the genetic parameters reflecting the genetic organization of cultivar Khachabi populations in Lebanon compare excellently with those of close wild relatives.

The introduction of clonal propagation through grafting techniques brings the advantages of clonal propagation, particularly the fixation of elite genotypes, as well as other genetic effects, and first the reduction of genotypic diversity and a marked differentiation among populations. These effects are not as rapid and dramatic as could be expected, because of the long generation time that maintains old genetic patterns, and because farmers adapt their practices

progressively, introducing vegetative reproduction in sexually propagated cultivars, as well as sexual recombination in the new clonal cultivars, thus generating new genetic diversity. As the evolution of farmers' practices gives more and more prevalence to vegetative reproduction, newly introduced populations of the clonal cultivar evolve differently, with a stronger loss of genetic diversity and less interactions with ancient cultivars. The observed evolution of the cultivar Halwani follows clearly such a pattern.

The almond self-incompatibility system plays a particular role in these developments, imposing outcrossing in traditional, sexually propagated, cultivars, favoring recombination among ancient cultivars and new introductions, but also imposing a threshold of minimal genetic diversity to avoid mate limitation and the attendant yield reductions. Beyond this threshold, the system would evolve towards the exclusive use of vegetative propagation, as in western horticulture, implying the adoption of the complete technical package of modern horticulture (establishment of polyclonal orchards or use of specialized pollinator cultivars, unless a truly self-compatible cultivar is developed), while resulting in a dramatic erosion of genetic diversity.

Conclusion

The population approach has allowed us to identify the effect of practices of farmers and nurserymen on the dynamics of genetic diversity within and among cultivars. Our results have shown different patterns of genetic diversity and population genetic structure for almond cultivars cultivated in Lebanon. Our study does not only contrast the effects of seed- and graft- propagation on genetic and genotypic diversity; it also shows that these propagation modes are not exclusive: farmers introduce clonal propagation in the seed-propagated cultivar while they maintain a diversity of genotypes within populations that are mostly graft-

propagated. In such a mixed system, genetic diversity is not negatively affected and orchard productivity is not limited by the availability of compatible pollinators.

Our study also detected long-distance exchanges of planting materials. Elucidating the social networks driving them will lead to a better understanding of the spatial and temporal dynamics of the genetic structure of the Lebanese almond and of the diversification processes at work. It will provide more in-depth information to develop a strategy for *in situ* conservation of cultivars and to reduce gene flow from introduced material to ancient orchards, as well as to natural populations of various *Prunus* species.

List of abbreviations

SSR: Simple Sequence Repeat

PCR: Polymerase Chain Reaction

MLGs: Multi Locus Genotypes

N: the number of sampled individuals

G: the number of distinct sampled genotypes

PIC: Polymorphic Information Content

R: genotypic richness

PI: Probability of Identity

LD: Linkage Disequilibrium

HWE: Hardy-weinberg Equilibrium

UHe: Unbiased expected heterozygosity

Ho: Observed heterozygosity

Fis: Wright's fixation index

Fst: genetic differentiation index

Dc: genetic distance

ENA: Excluding Null Alleles

INA: Including Null Alleles

IBD: Isolation By Distance

DAPC: Discriminant Analysis of Principal Components

MCMC: Markov Chain Monte Carlo

K: number of clusters

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and/or analyzed during the current study are available in the figshare repository, <https://dx.doi.org/10.6084/m9.figshare.5395459>.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was funded by the French-Lebanese *Program CEDRE 09 E F6/L 24*, the *Agropolis Foundation* project Fruit Med and Cirad core-found. This is a LIA O-LiFE contribution, number SA 11-2014.

Authors' contributions

BH collected and analyzed study materials, interpreted data as well as elaborated manuscript. LC coordinated the study and drafted the first version of manuscript. GC contributed in discussion of the analyses and in writing. LB contributed to the technical development of the genetic makers and to the acquisition of data; she took part in data analysis and interpretation and drafted the methods part of the manuscript. HIJ was a major contributor in writing the manuscript and provided guidance on the whole study. All authors read and approved the final manuscript.

Acknowledgements

The authors are indebted to all Lebanese farmers who contributed precious information on their germplasm and propagation practices. The authors are grateful to the CEFE “Service Commun de Marqueurs Génétiques en Ecologie” and the LabEx CeMEB “plate-forme séquençage-génotypage”, University of Montpellier II where genotyping was performed. The authors are grateful to Dr Michel Afram, President Director General of the Lebanese Agricultural Research Institute (LARI) for his support to this work. The authors also thank Mr Ahmad El Bitar and Mr Ali Chehade from the Biotechnology Department at LARI for their technical support at the first stage of field work and Dr Hafid Ahtak for his valuable introduction to data analysis. We thank Dr Marco Pautasso, Dr Philippe Jarne for their helpful comments on an early version of the manuscript.

References

1. Williams GC. Sex and Evolution. 3rd ed. Oxford: Princeton University Press; 1975.
2. Crow JF. Advantages of sexual reproduction. *Developmental Genetics* 1994; 5:205-213.
3. Roze D. Disentangling the Benefits of Sex. *Plos Biol* 2012;10: e1001321.
4. Austerlitz F, Mariette S, Machon N, Gouyon PH, Godelle B. Effects of colonization processes on genetic diversity: differences between annual plants and tree species. *Genetics* 2000; 154:1309–1321.

5. Hamrick JL, Godt MJW, Sherman-Broyles SL. Factors influencing levels of genetic diversity in woody plant species. *New Forest* 1992; 6:95–124.
6. Hamrick JL, Godt MJ. Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society B. Biological Sciences* 1996; 351:1291–1298.
7. Balloux F, Lehmann L, De Meeusb T. The Population Genetics of Clonal and Partially Clonal Diploids. *Genetics* 2003; 164:1635–1644.
8. Stoeckel S, Grange J, Fernandez-Manjarres JF, Bielger I, Franscaria-Lacoste N, Mariette S. Heterozygote excess in a self-incompatible and partially clonal forest tree species *Prunus avium* L. *Mol Ecol* 2006; 15:2109–2118.
9. Zohary D. Modes of evolution of plants under domestication. in W. F. Grant, ed. *Plant Biosystematics*. Toronto: Academic; 1984. p. 579–596.
10. Miller AJ, Gross BL. From forest to field: perennial fruit crop domestication. *Am J Bot* 2011; 98:1389–1414.
11. McKey D, Elias M, Pujol B, Duputie A. The evolutionary ecology of clonally propagated domesticated plants. *New Phytol* 2010; 186:318–332.
12. Zeder MA. *Documenting Domestication: New Genetic and Archaeological Paradigms*. Berkeley, California: University of California Press; 2006.
13. Browicz K, Zohary D. The genus *Amygdalus* L. (Rosaceae): species relationships, distribution and evolution under domestication. *Genet Resour Crop Ev* 1996; 43:229–247.
14. Ladizinsky G. On the origin of almond. *Genet Resour Crop Ev* 1999; 46:143–147.
15. Zohary D, Hopf M. *Domestication of plants in the old world*. 3rd ed. Oxford University; 2000.
16. Grasselly C, Crossa-Raynaud P. *L'amandier. Techniques agricoles et productions méditerranéennes*. Paris, XII: GP Maisonneuve et Larose; 1980.
17. Socias i Company R, Alonso JM, Kodad O, Gradziel TM. Almond. in Badenes ML, Byrne D, ed. *Fruit breeding*. Heidelberg, Germany: Springer verlag 2012; p 97–728.
18. Chalak L, Chehade A, Kadri A, Cosson P, Zanetto A, Dirlewanger E, Laigret F. Preliminary characterization of cultivated almonds (*Prunus dulcis* L.) in Lebanon by morphological traits and microsatellite markers. *Biol Tunis* 2006; 4:5–11.

19. Fernández i Marti, ÁF, Alonso JM, Espiau MT, Rubio-Cabetas MJ, Socias i Company R. Genetic diversity in Spanish and foreign almond germplasm assessed by molecular characterization with simple sequence repeats. *J Am Soc Hortic Sci* 2009; 134: 535–542.
20. El Hamzaoui A, Oukabli A, Charafi J, Moumni M. Assessment of genetic diversity of Moroccan cultivated almond (*Prunus dulcis* Mill.D. A. Webb) in its area of extreme diffusion, using nuclear microsatellites. *Am J Plant Sci* 2012; 1294–1303.
21. Bourguiba H, Khadari B, Krichen L, Trifi-Farah N, Santoni S, Audergon JM. Grafting versus seed propagated apricot populations: two main gene pools in Tunisia evidenced by SSR markers and model-based Bayesian clustering. *Genetica* 2010; 138:1023–1032.
22. Gouta H, Ksia E, Buhner T, Moreno MA, Zarrouk M, Mliki A, Gogorcena Y. Assessment of genetic diversity and relatedness among Tunisian almond germplasm using SSR markers. *Hereditas* 2010; 147:283–292.
23. Khadari B, Charafi J, Moukhli A, Ater M. Substantial genetic diversity in cultivated Moroccan olive despite a single major cultivar: a paradoxical situation evidenced by the use of SSR loci. *Tree Genet Genomes* 2008; 4:213–221.
24. Beghè D, Ferrarini A, Ganino T, Fabbriet A. Molecular characterization and identification of a group of local *Olea europaea* L. varieties. *Tree Genet Genomes* 2011; 7:1185–1198.
25. Achtak H, Ater M, Oukabli A, Santoni S, Kjellberg F, Khadari B. Traditional agroecosystems as conservatories and incubators of cultivated plant varietal diversity: the case of fig (*Ficus carica* L.) in Morocco. *BMC Plant Biol* 2010; 10:28.
26. Delplancke M, Alvarez N, Espindola A, Joly HI, Benoit L, Brouck E, Arrigo N. Gene flow among wild and domesticated almond species: insights from chloroplast and nuclear markers. *Evol Appl* 2012; 5:317–329.
27. Wang Z, Kang M, Liu H, Gao J, Zhang Z, et al. High-Level Genetic Diversity and Complex Population Structure of Siberian Apricot (*Prunus sibirica* L.) in China as Revealed by Nuclear SSR Markers. *PLoS ONE* 2014; 9(2): e87381. doi:10.1371/journal.pone.0087381
28. Post GE, Dinsmore JE. *Flora of Syria, Palestine, and Sinai*. 2nd ed. American university of Beirut: Publications of the Faculty of Arts and Sciences, 1932; 1:658.

29. Mouterde P. Nouvelle Flore du Liban et de la Syrie. Editions de l'imprimerie catholique, Beirut, Lebanon 1966; 1:563.
30. Talhouk SN, Lubani RT, Parmaksizian LS, Nehme GA. Survey and characterization of almond germplasm in Lebanon. *Acta Hort* 1997; 470:57–65.
31. Talhouk SN, Lubani RT, Baalbaki R, Zurayk R, Al Khatib A, Parmaksizian L, Jaradat AA. Phenotypic diversity and morphological characterization of *Amygdalus* species in Lebanon. *Genet Resour Crop Evol* 2000; 47:93–104.
32. Delplancke M, Alvarez N, Benoit L, Espindola A, Joly HI, Neuenschwander S, Arrigo N. Evolutionary history of almond tree domestication in the Mediterranean basin. *Mol Ecol* 2013; 22:1092–1104.
33. Chalak L, Chehade A, Kadri A. Morphological characterization of cultivated almonds in Lebanon. *Fruits* 2007; 62:177–186.
34. Cipriani G, Lot G, Huang WG, Marrazzo MT, Peterlunger E, Testolin R. AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* L. Batsch]: isolation, characterisation and cross-species amplification in *Prunus*. *Theo Appl Genet* 1999; 99:65–72.
35. Testolin R, Marrazzo T, Cipriani G., Quarta R, Verde I, Dettori MT, Pancaldi M. Microsatellite DNA in peach (*Prunus persica* L. Batsch) and its use in fingerprinting and testing the genetic origin of cultivars. *Genome* 2000; 43:512–520.
36. Sosinski B, Gannavarapu M, Hager LD, Beck LE, King GJ, et al. Characterization of microsatellite markers in peach (*Prunus persica* L. Batsch). *Theo Appl Genet* 2000; 101:421–428.
37. Dirlewanger E, Cosson P, Tavaud M, Aranzana MJ, Poizat C, Zanetto A, Arus P. Development of microsatellite markers in peach [*Prunus persica* L. Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). *Theo Appl Genet* 2002; 105:127–138.
38. Aranzana MJ, Garcia-Mas J, Carbo J, Arus P. Development and variability analysis of microsatellite markers in peach. *Plant Breeding* 2002; 121:87–92.
39. Arnaud-Haond S, Belkhir K. GENCLONE: a computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. *Mol Ecol Notes* 2007; 7:15–17.

40. Peakall R, Smouse PE. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 2006; 6:288–29.
41. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 1980; 32:314–331.
42. Marshall TC, Slate J, Kruuk LEB, Pemberton JM. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol Ecol* 1998; 7:639–655.
43. Dorken ME, Eckert CG. Severely reduced sexual reproduction in northern populations of a clonal plant, *Decodon verticillatus* (Lythraceae). *J Ecol* 2001; 89:339–350.
44. Raymond M, Rousset F. GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *J Hered* 1995; 86:248–249.
45. Rousset F. GENEPOP'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Res* 2008; 8:103–106.
46. Dempster AP, Laird NM, Rubin DB. Maximum Likelihood from incomplete data via the *EM* algorithm (with discussion). *J R Stat Soc Series B* 1977; 39:1–38.
47. Szpiech ZA, Jakobsson M, Rosenberg NA. ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics* 2008; 24:2498–2504.
48. Wright S. The genetical structure of populations. *Ann. Eugethes* 1951; 15:323–54.
49. Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F. GENETIX v.4.05 logiciel sous Windows pour la génétique des populations. <http://kimura.univ-montp2.fr/genetix/> (2004). Accessed 21 Sept 2014.
50. Chapuis MP, Estoup A. Microsatellite Null Alleles and Estimation of Population Differentiation. *Mol Biol Evol* 2007; 24:621–631.
51. Cavalli-Sforza LL, Edwards AWF. Phylogenetic analysis: models and estimation procedures. *Evolution* 1967; 21:550–570.
52. Mantel N. The detection of disease clustering and generalised regression approach. *Cancer Res* 1967; 27:209–220.
53. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics* 2000; 155:945–959.

-
54. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 2005; 14:2611–2620.
 55. Earl DA, Bridgett VHM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Res* 2012; 4:359–361.
 56. Jakobsson M, Rosenberg NA. CLUMPP: cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 2007; 23:1801–1806.
 57. Rosenberg NA. DISTRUCT: a program for the graphical display of population structure. *Mol Ecol Notes* 2004; 4:137–138.
 58. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics* 2010; 11:94.
 59. Jombart T. Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* 2008; 24:1403–1405.
 60. R Foundation for Statistical Computing. R: A Language and Environment for Statistical Computing, Reference Index Version 3.1.2. R Foundation for Statistical Computing, Vienna, Austria. 2014. Accessed 10 Dec 2014.
 61. Legendre P, Legendre L. *Numerical Ecology*. 2nd English ed. Elsevier, Amsterdam; 1998.
 62. Campoy JA, Lerigoleur Balsemin E, Christmann H, Beauvieux R, Girollet N, Quero-Garcia J, Dirlewanger E, Barreneche T. Genetic diversity, linkage disequilibrium, population structure and construction of a core collection of *Prunus avium* L. landraces and bred cultivars. *BMC Plant Biol* 2016; 16 (1): 15. DOI : 10.1186/s12870-016-0712-9
 63. Bourguiba H., Audergon JM, Krichen L, Trifi-Farah N, Mamouni A, Trabelsi S, D’Onofrio C, Asma BM, Santoni S, Khadari B. Loss of genetic diversity as a signature of apricot domestication and diffusion into the Mediterranean Basin. *BMC Plant Biol* 2012; 12:49. doi:10.1186/1471-2229-12-49
 64. Chybicki IJ, Burczyk J. Simultaneous Estimation of Null Alleles and Inbreeding Coefficients. *J Hered* 2009; 100:106–113.

65. Elias M, Santos Muhlen G, McKey D, Roa AC, Tohme J. Genetic diversity of traditional south American landraces of cassava (*Manihot esculena Crantz*): An analysis using microsatellites. *Econ Bot* 2004; 58:242–256.
66. Batnini M, Krichen L, Bourguiba H. Trifi-Farah N, Ruiz D, Martinez-Gomez P, Rubio M. Comparative analysis of traditional and modern apricot breeding programs: A case of study with Spanish and Tunisian apricot breeding germplasm. *Span J Agric Res*. 2016. 14. e0706.
67. Pritchard JK, Wen X, Falush D. Documentation for STRUCTURE software: Version 2.3. <http://pritch.bsd.uchicago.edu/structure.html> (2010). Accessed 11 Feb 2011.
68. Jackson JF. Gene flow in pollen in commercial almond orchards. *Sex Plant Reprod* 1996; 9:367–369.

Additional files

Additional file 1: Figure S1.tif The geographic locations of the collected populations.

A total of 14 ‘Khachabi’ and 11 ‘Halwani’ populations were sampled, covering the four major agro-climatic zones.

Additional file 2: Table S1.xls Statistics associated with microsatellite loci.

Na, total number of alleles; PIC, polymorphic information content; Ho, observed heterozygosity; He, unbiased expected heterozygosity; Fis, inbreeding index calculated based on a matrix of N individuals or G genotypes; f (Null), estimation of significant frequency of null alleles. (a) Cipriani *et al.* 1999; (b) Testolin *et al.* 2000; (c) Sosinski *et al.* 2000; (d) Dirlewanger *et al.* 2002; (f) Aranzana *et al.* 2002. * HW significance with Bonferroni correction ($P < 0.01$). Fis_(N) values were calculated with all trees found including repeated clone genotype. Fis_(G) values were calculated with only one individual per multi locus genotype (MLG).

Additional file 3: Table S2.xls Genotypic linkage disequilibrium for pairwise comparisons between loci conducted over different data sets without redundant MLGs.

– data not available; The coefficient of correlation is indicated for $P > 0.05$.

Additional file 4: Figure S2.tif Mean number of alleles per locus as a function of sample size.

N, number of MLGs. Diamonds for global dataset, squares for ‘Khachabi’, triangles for ‘Halwani’.

Additional file 5: Table S3.xls Confidence intervals for Fis values at 95%.

Fis calculated for sample size over 15 individuals (N) or genotypes (G).

Additional file 6: Figure S3 Plot of genetic distance (Dc) and geographic distance for A) 14 ‘Khachabi’ populations; B) 11 ‘Halwani’ populations. Significance at $\alpha = 0.05$.

Additional files

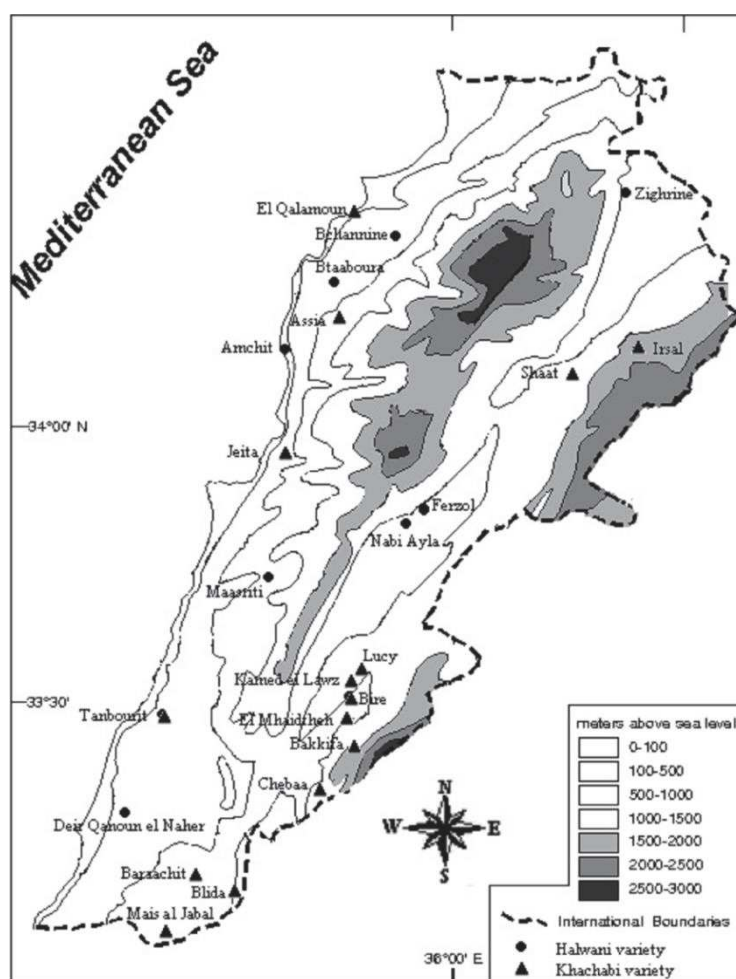


Fig.S1 The geographical locations of the collected populations. 14 ‘Khachabi’ and 11 ‘Halwani’ populations were sampled, covering the four major agro-climatic zones.

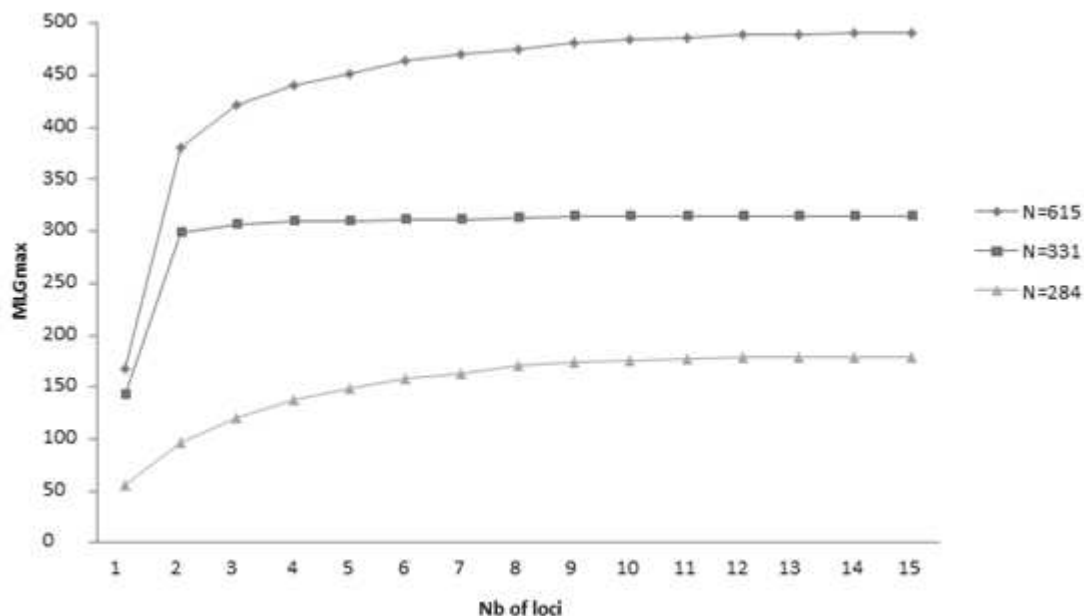


Fig. S2 Mean number of alleles per locus as a function of sample size (N; number of MLGs).

Losange for overall dataset. Square for Khachabi. Triangle for Halwani.

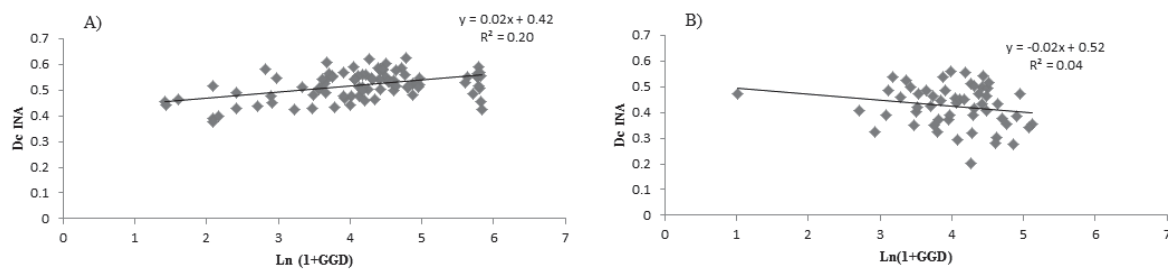


Fig. S3 Plot of genetic distance (Dc) and geographic distance for A) 14 'Khachabi' populations; B) 11 'Halwani' populations. Significance at $\alpha = 0.05$.

Les tableaux supplémentaires seront disponibles sur le site du journal BMC Plant biology

Discussion générale et perspectives

Dynamique de la diversité génétique et diffusion des amandiers cultivés au Liban, influence des savoirs et des pratiques: regards croisés génétique – anthropologie

Nos travaux confirment que l'amandier est un modèle intéressant pour appréhender l'influence de l'Homme sur la structuration et la dynamique de la diversité des plantes cultivées. Dans ma thèse, j'ai opté pour une approche interdisciplinaire qui est basée sur la génétique des populations et l'anthropologie pour comprendre la dynamique des processus de diversification variétale de l'amandier au Liban. Je me suis appuyée sur la littérature historique et archéobotanique pour mettre en situation les différentes étapes de l'impact de l'homme sur les forces évolutives (sélection et migration) à différentes échelles spatiales et temporelles. Ensuite j'ai évalué la diversité intra-spécifique culturelle en analysant le système de nomenclature et la classification locale. Finalement, j'ai évalué l'impact des différentes pratiques de propagation et de diffusion sur la diversité génétique à l'échelle des vergers.

L'originalité de cette thèse est aussi dans l'utilisation d'une approche de génétique des populations qui a dévoilé l'importance des flux de gènes inter et intra variétale dans le processus de diversification variétale de l'amandier aussi bien au niveau du Bassin Méditerranéen qu'au niveau du Liban et des vergers. D'autre part, les résultats de cette thèse montrent que la diversité de l'amandier cultivé est restée longtemps dans un équilibre dynamique alors que la dynamique actuelle du processus évolutif est limitée par l'ampleur du greffage. Un complément intéressant pour poursuivre cette étude serait d'évaluer la diversité des espèces spontanées apparentées et leurs usages, passés et présents, afin de mieux mesurer leur importance dans le pool génétique de l'amandier cultivé aujourd'hui.

Cette thèse a mis en évidence l'interaction entre les pratiques de reconduction des vergers et la diversité inter et intra variétale de l'amandier cultivé au Liban. Nous avons montré que les connaissances des acteurs et leurs pratiques culturelles influencent la diversité à l'échelle du verger. Cette diversité repose sur les pratiques de reconduction des vergers et d'échanges

locaux qui ont conduit à une diversité culturelle autour de l'amandier cultivé. L'introduction de nouveaux cultivars depuis le siècle dernier et l'adoption du mode de propagation végétatif par greffage ont été accompagnés d'un système dynamique de dénomination. Les différentes pratiques de sélection locale et d'introduction de nouveaux cultivars sont liés au goût pour la diversité et pour l'expérimentation dont font preuve les agriculteurs, ce qui contribue à augmenter la diversité variétale. Cependant, une pression du milieu importante (sécheresse, ravageurs, etc.) peut conduire à une perte de génotypes, voire de cultivars à l'échelle du verger et même à l'échelle du pays.

1) Dynamique d'introduction, diversité nommée et diversité variétale

Dans le Chapitre 1, nous avons analysé la diversité génétique et la structuration de onze cultivars. Nos résultats montrent que les arbres portant un même nom, donné par les informateurs, sont généralement dans le même pool génétique. Cependant, une discordance entre certains noms donnés et la structuration génétique a été observée. De plus, nos résultats indiquent que la diversité nommée et la diversité variétale ont évolué avec le temps.

Notre analyse de la diversité génétique montre que le cultivar le plus anciennement cultivé, nommé 'Khachabi', est différencié en deux pools de gènes structurés géographiquement, cependant la diversité nommée à l'échelle intra-variétale ne correspond pas à ces deux unités génétiques. La discordance entre la diversité nommée et la diversité génétique est attendue à ce niveau, car la nomenclature locale est basée sur le simple critère morphologique de la taille du fruit.

L'ancien cultivar introduit nommé 'Ferk'/'Halwani' est aussi différencié génétiquement en deux pools de gènes. Les résultats du Chapitre 3 montrent que cette différenciation génétique existe même entre des individus présents dans le même verger. Les noms 'Ferk' et 'Halwani' sont présents dans les deux pools de gènes identifiés dans le Chapitre 1 et confirmés dans le

Chapitre 3. La diversité nommée est due uniquement à la dynamique temporelle du système de nomenclature. En effet, ‘Ferk helou’ (coque tendre et amandon doux) identifié dans les enquêtes du service du Général Gouraud en 1921 est probablement devenu ‘Halwani’ après l’introduction d’autres cultivars pour le différencier de ces cultivars portant les mêmes caractéristiques ; ces cultivars nouvellement introduits sont identifiés dans la taxonomie locale par le nom d’origine du cultivar ou celui du pays d’où il vient. Cette discordance entre la diversité nommée et la diversité génétique est alors liée au processus de nomenclature et à l’existence de catégories englobantes. Par ailleurs, la diversité nommée intra-variétale de ‘Halwani’, ‘Halwani lebb’ (amandon simple) et ‘Halwani lebbein’ (amandon double), se base sur un critère morphologique. Ce critère quantitatif soumis à la sélection est donc impossible à différencier avec les marqueurs neutres que nous avons utilisés.

Les résultats de l’analyse de structuration génétique issus du Chapitre 1 montrent la présence d’un pool de gènes hétérogène incluant les types nommés ‘Halwani’, ‘Awja’, ‘Bandouk Awja’ et ‘Itali’. Cette discordance entre la diversité variétale et la diversité nommée est probablement liée aux pratiques de sélection locale et de propagation sexuée adoptées pour la diversification de l’amandier qui sont indépendantes du processus de nomenclature. Ainsi, la reproduction sexuée favorise la production d’hybrides entre les cultivars; dans certains cas ces hybrides peuvent être multipliés végétativement et puis nommés. La distance génétique entre ces différents cultivars est donc réduite et le nombre de types nommés augmente, expliquant la présence de différents types nommés dans un même pool de gènes. La discordance peut aussi être due à des reprises du porte-greffe aux dépens du greffon comme on l’a observé entre ‘Halwani’ et ‘Khachabi’ dans le Chapitre 3.

Le type nommé ‘Nahali’ forme une unité génétiquement différenciée des autres types nommés. Cet accord entre la diversité nommée et la diversité variétale, un type nommé équivalent à un cultivar, est due à la propagation végétative de ce cultivar présent uniquement

dans la région de Batroun. De plus, bien qu'il soit assez important dans la région du Liban Nord, le type nommé 'Oum Omar' n'a pas pu être pris en compte dans nos analyses génétiques; en effet, les dégâts sur les feuilles de ce cultivar, endommagées par le phytoplasme, sont très importants dans cette région (Choueiri, et al., 2003).

Par ailleurs, notre étude a montré aussi une discordance entre la dénomination donnée par les sélectionneurs et la diversité variétale nommée par les informateurs. Les cultivars d'amandier les plus récents provenant de la France ont gardé leurs noms d'origine 'Ferragnès', 'Ferraduel' et 'Lauranne'. Ces trois cultivars, appartiennent au même pool de gènes. Ceci s'explique par le fait que 'Ferragnès' et 'Ferraduel' sont issues d'un même croisement (Grasselly and Crossa-Raynaud, 1980), et que 'Lauranne' est le descendant d'un croisement avec 'Ferragnès' (Grasselly and Dirlewanger, 2016). Cependant ces noms n'ont pas été cités dans les listes libres; ces cultivars peuvent correspondre aux types nommés localement 'Faransi Abrach', ou 'Faransi Mallisi'.

Quant à l'amandier amer, localement nommé 'Barri', c.a.d. « sauvage », il est inclus dans le système taxonomique local de l'amandier cultivé (Chapitre 2). Cependant, nous avons analysé la diversité génétique de l'amandier cultivé sans avoir pris en compte l'amandier amer ce qui ne nous permet pas de vérifier si celui-ci constitue une entité génétique à part. Par ailleurs, Delplancke et al. (2013) ont démontré la contribution de *P. orientalis*, apparenté spontané de l'amandier cultivé présent au Liban, dans la diversification de l'amandier cultivé; nous n'avons pas d'information qui nous permette de connaître la place de cette espèce dans la taxonomie locale.

Les agriculteurs intègrent l'amandier amer dans le système de propagation de l'amandier cultivé en l'utilisant en tant que porte-greffe. Comme les populations d'amandier amer sont proches des vergers, il serait intéressant d'évaluer les flux de gènes et de quantifier

l'introgression entre ces deux compartiments. Une dynamique d'introgression a été mise en évidence entre l'olivier issus de semis (oléastre) et l'olivier cultivé au Maroc (Haouane, 2012). L'analyse des pratiques paysannes a montré que l'oléastre est, comme l'amandier amer, utilisé comme porte-greffe et de plus il est aussi sélectionné pour la production de l'huile.

2) Dynamique de la diversité génétique, diffusion et pratiques de propagation

Dans cette partie, les résultats des analyses de la diversité génétique effectuées dans les Chapitres 1 et 3 sont discutés pour montrer comment la dynamique de la diversité variétale est soumise à des facteurs sociaux, à différentes échelles spatio-temporelles ; cette réflexion s'appuie sur les résultats du Chapitre 2 et sur quelques enquêtes complémentaires effectuées auprès des pépiniéristes.

Ainsi, le cultivar 'Khachabi' appartient à deux pools de gènes, actuellement distribués l'un du Liban à la Turquie incluant la Syrie et l'autre allant du Liban à l'Italie ; ceci peut être interprété comme le résultat d'une diffusion progressive par graines à l'échelle de l'ensemble du Bassin Méditerranéen depuis les grands voyages *circum* Méditerranée. A l'échelle du Liban, les vergers du cultivar 'Khachabi' sont encore abondants ; les enquêtes montrent que sa diffusion à cette échelle est locale, limitée le plus souvent au village. Cette dynamique locale explique le pattern d'isolement par la distance caractéristique des vergers de 'Khachabi' présents dans la plaine de la Bekaa (Chapitre 3). Le déficit en hétérozygotie observé dans les vergers de 'Khachabi' peut être expliqué par l'endogamie entre les arbres à cette l'échelle. Une diversité d'allèles d'auto-incompatibilité doit donc nécessairement être présente dans chaque verger pour la fructification et pour le maintien de la population. Différentes études ont recensé, jusqu'à présent, une cinquantaine d'allèles d'auto-incompatibilité (Kodad et al., 2010 ; Hafizi et al., 2013) mais aucune n'a eu une approche de génétique des populations. En perspective, il serait intéressant d'évaluer la diversité des

allèles d'auto-incompatibilité présente chez l'amandier à l'échelle intra-variétale et surtout à l'échelle des vergers de 'Khachabi' ainsi que leur distribution spatiale à l'échelle du Liban.

Les cultivars introduits au Liban ne sont pas issus des pools de gènes méditerranéens étudiés, à l'exception des cultivars français (Ferragnès, Ferraduel et Lauranne) introduits récemment au Liban. Les analyses génétiques mettent en évidence une diffusion de clones à l'échelle du Bassin méditerranéen. A l'échelle du pays, ces analyses ont aussi mis en évidence des géotypes communs entre les différents vergers, reflétant une dynamique de diffusion des clones. Les enquêtes menées auprès des informateurs responsables de ces vergers dévoilent une diversité de provenances des plants greffés ou des greffons, basée le plus souvent sur du matériel venant des pépinières et occasionnellement sur les réseaux d'échange entre agriculteurs et rarement sur les ONGs et les services agricoles de l'état (Annexe 2).

Cette approche pluridisciplinaire a mis en évidence l'importance des pratiques des agriculteurs qui agissent directement sur la diversité en introduisant ou en éliminant des cultivars. D'une part, l'origine des plants greffés et des greffons, données par les agriculteurs enquêtés, a été confirmée par la distribution géographique des géotypes communs entre les vergers. D'autre part, la comparaison entre les deux cultivars phares 'Khachabi' et 'Halwani' a permis de mettre en évidence l'impact des pratiques de propagation sur la diversité génétique. Une réduction de la diversité a été observée à l'échelle du cultivar et des vergers de 'Halwani' à propagation végétative comparée à celle du cultivar 'Khachabi' qui est propagé surtout par semis. La distribution spatiale des clones de ces deux cultivars à travers le Liban a montré la présence de certains géotypes communs entre différents vergers, présents dans différentes zone éco-géographiques (clones communs); un réseau d'échange des cultivars est donc confirmé par la présence de ces clones communs. Certains clones sont cependant présents dans un seul verger (clones privés), ce qui souligne les pratiques de multiplication

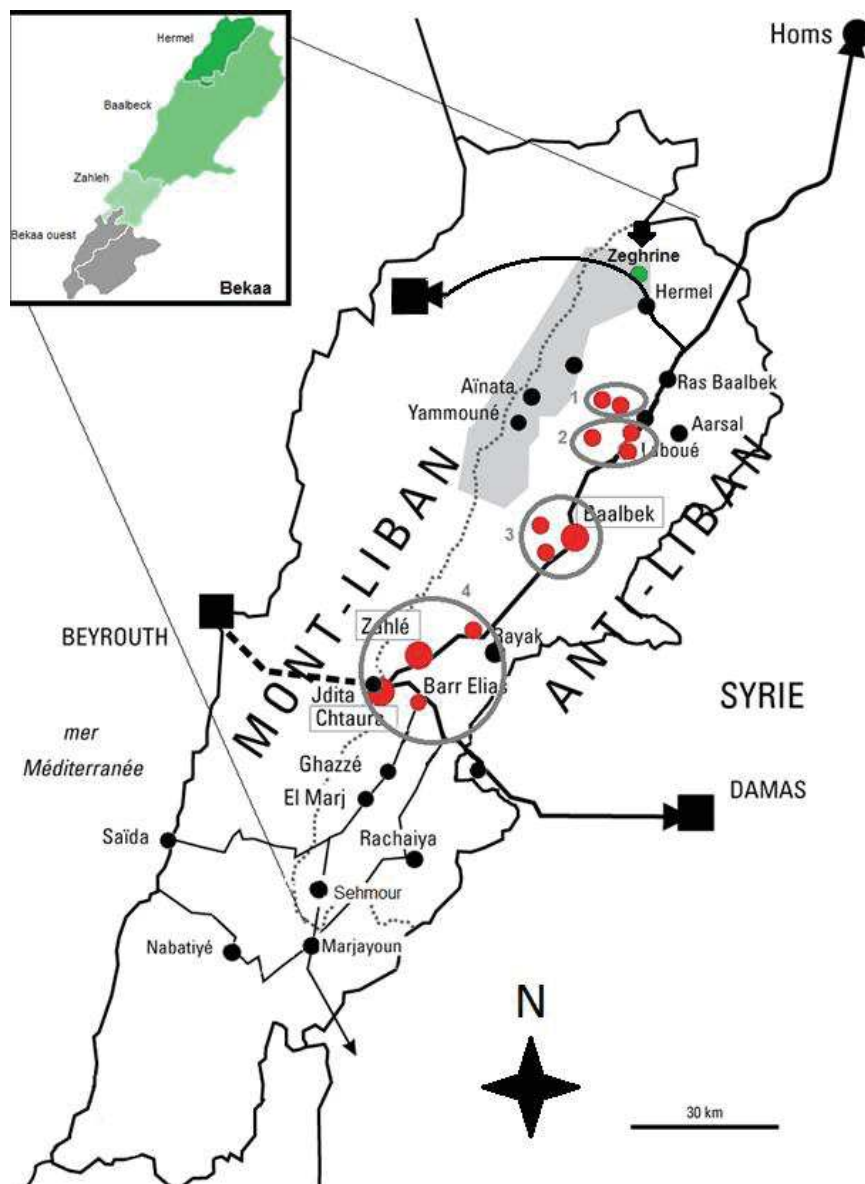


Figure IV-1 Carte du Liban, distribution des pépiniéristes enquêtés.

La flèche indique le village de Zeghrine, où on a enquêté les agriculteurs sur l'origine de cultivars d'amandier. Les Pépiniéristes sont représentés par des points regroupés dans des cercles numérotés. Chaque numéro correspond à un groupe de pépiniéristes. Les autres points correspondent aux principaux acheteurs ou clients.

végétative des agriculteurs et la propagation vraisemblable de leur propre matériel végétal.

Des enquêtes complémentaires, réalisées auprès des agriculteurs de Zeghrine sur les questions d'origine du matériel qu'ils cultivent, ont permis de mettre en évidence le rôle des pépiniéristes. Cela nous a amené à réaliser quelques enquêtes préliminaires auprès de pépiniéristes dans la plaine de la Bekaa. A partir des suggestions des agriculteurs de Zeghrine, nous avons enquêté, par une approche « boule de neige » (Heckathorn, 1997), 20 pépiniéristes depuis le village de Jabbouleh jusqu'à la zone de Zahlé (voir carte ci-contre, Figure IV-1). Trois types de pépiniéristes peuvent être identifiés dont les comportements sont différents. Les premiers sont des agriculteurs qui complètent leurs activités de production agricole par la production de plants. Ils multiplient essentiellement des arbres fruitiers et leurs clients sont des agriculteurs des villages montagneux aux alentours ; l'origine première du matériel qu'ils multiplient semble être le verger d'un ancien salarié du Ministère de l'Agriculture qui n'a pas pu nous donner plus d'information lui-même sur l'origine du matériel de son verger. La seconde catégorie de pépiniéristes, autour de la ville de Zahlé, regroupe des professionnels qui gèrent leurs pépinières de père en fils ; ils produisent en général non seulement des arbres fruitiers mais aussi des arbres forestiers et des fleurs ; ces pépiniéristes s'approvisionnent parfois en retournant dans les vergers d'anciens clients agriculteurs pour prélever des greffons. Enfin, entre les deux zones, autour de Baalbeck se trouvent surtout des pépiniéristes-commerçants ; ceux-ci ne produisent pas de plants mais s'approvisionnent soit chez les pépiniéristes de Zahlé, soit dans le Nord du Liban, autour de Tripoli et ne sont pas spécialisés sur les arbres fruitiers ; l'amandier ne représente que 15% de leurs ventes.

Ce travail permet de mettre en évidence, parmi les clones hybrides entre 'Kachabi' et 'Halwani', la présence de matériel provenant de pépinières ; les pépiniéristes pratiquent donc aussi vraisemblablement une reproduction combinée du semis et du greffage dans la

propagation des plants qui ne sont pas tous du matériel certifié. D'ailleurs, les pépiniéristes sont les seuls à reconnaître trois catégories de cultivars selon leur ancienneté: les arbres du cultivar Khachabi présents dans les anciens vergers avant l'introduction de cultivars améliorés; les premiers cultivars greffés, notamment 'Halwani', introduits dans les années 60; ensuite, les nouveaux cultivars améliorés et certifiés introduits depuis l'année 2000. Cette dernière phase d'introduction a coïncidé avec la création de la Coopérative Nationale des Pépiniéristes « Mashatel Lebnen » et la mise en place d'un catalogue de variétés disponibles. Seuls les pépiniéristes ont clairement parlé de ces trois périodes, les agriculteurs ne les ont pas différenciées dans leurs discours. Cependant l'analyse de la taxonomie locale, proposée dans le Chapitre 2, reflète une dynamique temporelle de dénomination et d'usage des critères de classification.

Une étude plus fine des pratiques de sélection et de propagation des pépiniéristes au Liban et de leurs connaissances devrait permettre de mieux comprendre l'impact de ces acteurs qui sont probablement devenus des acteurs clés de voute pour la diversité de l'amandier cultivé et sa dynamique. De plus, ces pépiniéristes sont organisés en coopérative, un réseau qui facilite l'échange et la diffusion de matériel végétal.

3) Quelques implications pour la conservation

« La préservation des ressources n'est pas une fin en soi, mais le moyen d'assurer que les ressources génétiques animales et végétales restent à la disposition des générations actuelles et futures. Après avoir identifié et caractérisé ces ressources, on peut adopter, pour les préserver, soit l'approche in situ, soit l'approche ex situ. Ces deux approches ne sont pas incompatibles: elles peuvent se compléter, ce qui atténue les inconvénients de l'une ou l'autre méthode. Si l'on excepte un petit nombre de variétés ou d'espèces qui sont largement utilisées, particulièrement dans les pays développés l'expérience prouve que, c'est seulement en variant les stratégies de préservation que l'on pourra assurer le maintien de la diversité. »

Valorisation la diversité de la nature, <http://www.fao.org/docrep/004/V1430F/V1430F06.htm>

Au Liban sous l'égide de la FAO, une stratégie nationale pour la conservation et la gestion des ressources phytogénétiques pour l'alimentation et l'agriculture a été récemment établie

A. Conservation Component

Specific objective 1. To ensure reliable on-farm and *in situ* conservation of PGRFA

This specific objective responds to the need to promote *in situ* conservation of cultivated PGRFA and their wild relatives taking into account the reciprocal relation existing between local communities and PGRFA. The establishment of biocultural territories is proposed, referring to the set of biological resources, comprising PGRFA diversity in terms of food and feed, the agricultural systems, and the landscapes and ecosystems which host them, as well as to the traditions and good agricultural practices of rural populations, including local communities, which are often the custodians of cultivated PGR and de-facto guardians of the agrobiodiversity. The establishment and appropriate management of biocultural territories require baseline information describing all the important aspects of such territories, as well as monitoring mechanisms. These territories are therefore considered integral parts of on-farm and *in situ* conservation strategies. Reinforcing and acknowledging local seed systems as fundamental elements for maintaining agrobiodiversity in biocultural territories is crucial for the effective conservation of cultivated PGRFA.

Specific objective 2. To ensure *ex situ* conservation of PGRFA and improve their access and usefulness

This specific objective is to improve the *ex situ* conservation of PGRFA and to ensure the access to germplasm and information between gene bank and institutions. A national programme for germplasm collection based on inventories, in consultation with experts and using GIS tools, with emphasis on either endangered materials or materials potentially adapted to climate change is proposed. The national genebank should guarantee the genetic integrity and phytosanitary quality / viability of germplasm conserved through adequate documentation, characterization and evaluation, applying agreed standards and best practices. A national repository for native lignified species is strongly required. The interesting germplasm should be available to users including materials for direct planting and research or plant breeding. The long-term conservation of non-orthodox germplasm, through collections of "live" plants in the field should be ensured. Where the context allows, multiplying germplasm for its distribution to farmers, particularly in term of seed shortage, as well as after natural disasters should be accomplished. A national network will be established through the NISM and GPA2 website linking data base with inventory and monitoring systems and strengthening the liaison between the stakeholders (national genebank, universities, local communities, civil society) and promoting cooperation among them, avoiding duplication of efforts. Access to and distribution of germplasm to users through the Multilateral System of the ITPGRFA or other relevant national, regional or international legislation should be promoted and facilitated.

Source : Chalak 2015.

impliquant la contribution de plusieurs institutions et organismes de recherche. Cette stratégie vise un engagement des acteurs dans tous les secteurs concernés pour atteindre les dix objectifs fixés par le Comité National pour conserver et valoriser la biodiversité ainsi qu'en assurer un usage durable (Chalak, 2015).

i) Conservation *ex situ*

La conservation *ex situ* constitue un objectif spécifique de la stratégie nationale des ressources phytogénétiques (cf. encadré). Une collection de base (arboretum) pour les arbres fruitiers à noyaux qui sont cultivés au Liban a été réalisée à l'IRAL à Tal Amara (Institut de Recherches Agronomique Libanais, institut gouvernemental sous la tutelle du Ministère de l'Agriculture).

Notre travail a montré une grande diversité intra et inter variétale des cultivars libanais ; il serait donc intéressant d'évaluer la diversité génétique des accessions d'amandier présentes dans la collection nationale afin d'évaluer si cette collection représente bien la diversité de l'amandier cultivé au Liban. Il sera aussi important de réaliser des collections complémentaires de base dans les différentes zones éco-géographiques du Liban afin de préserver la diversité régionale. L'ampleur de la diversité observée dans notre échantillonnage rendra cependant difficile l'établissement d'une "core collection" nationale et le choix des cultivars à préserver pour les collections régionales ; une réflexion approfondie sur la mise en place d'un protocole de conservation *in situ* nous semble donc nécessaire.

ii) Conservation *in situ*

La conservation *in situ* est aussi un des objectifs spécifiques de la stratégie nationale phytogénétiques (cf. encadré). La gestion à la ferme de la diversité génétique des cultures est une composante essentielle des efforts nationaux et internationaux visant à conserver les ressources phytogénétiques (FAO 2016). La réussite de la conservation *in situ* dépend des décisions des agriculteurs (Jarvis and Hodgkin, 2000). Selon Brush 1991, l'objectif de la

conservation *in situ* des plantes cultivées est d'encourager les agriculteurs à continuer de planter les différents cultivars locaux, en leurs donnant un surplus de valeur, morale et/ou monétaire (Marchenay, 2005).

Dans notre étude, l'évaluation de la dynamique de la diversité des variétés d'amandier cultivées au Liban par une approche interdisciplinaire permet de mieux cibler des zones qui peuvent faire l'objet de plan de conservation géré par les agriculteurs. Par exemple, les agriculteurs du village de Zeghrine ont conservé des savoirs et des savoir-faire anciens bien qu'ils aient opté pour une dynamique de diversification variétale qui se développe au détriment de la diversité des cultivars anciens. Cependant, ils préservent encore des arbres du cultivar traditionnel local 'Khachabi' en bordure de verger pour assurer une bonne pollinisation ; ils se servent aussi de la forme spontanée ou féral 'Kachabi morr' comme porte-greffe, y compris pour greffer des variétés améliorées introduites. En s'appuyant sur ces pratiques locales, une stratégie nationale pourrait être établie et mise en œuvre pour l'amandier.

Nos résultats montrent que l'enjeu de la conservation de la diversité intra-spécifique de l'amandier demeure dans la tension entre l'introduction de variétés améliorées et la propagation du matériel local. Le maintien, au moins partiel, de pratiques de semis et la multiplication clonale d'hybrides ainsi que les réseaux d'échange de matériel végétal entre agriculteurs pourraient être encouragés pour assurer un équilibre avec le développement de pépinières diffusant du matériel certifié et maintenir une dynamique évolutive de l'amandier.

Dans la région de la Bekaa ouest, nos enquêtes ont montré que l'introduction de nouveaux cultivars met en danger la persistance du cultivar traditionnel 'Khachabi' ; celui-ci représente cependant selon nos résultats un réservoir de diversité qui peut à son tour servir pour des programmes de sélection locale.

Specific objective 4. To enhance farm diversification and improve farmers livelihood

Introducing greater number of crops for climate change risk management and health, food and nutrition security is improved allowing the farms to produce an important amount of edible species, directly benefiting farmers and promoting new income-generating alternatives. Greater use of native local crops, landraces and varieties in existing food and agricultural assistance programmes is promoted. Local knowledge on PGR and underutilized edible species is documented, complemented with scientific information and promoted. Diversification of PGR value-added products should be prioritized, and transformation and commercialization processes should be promoted.

Specific objective 5. To enhance utilization of inter- and intraspecific diversity

This specific objective promotes plant breeding using participatory, conventional and molecular methods for developing varieties with characteristics that enable adaptation to climate change and meet the diverse needs and preferences of farmers. Both inter- and intra-specific diversity of crops and associated local knowledge are fully deployed in agricultural innovation processes. Producers' organizations and their entrepreneurial capacity for producing and selling quality seeds and agricultural products are strengthened.

Source : Chalak 2015.

iii) Conservation des ressources génétiques spontanées

Un complément très intéressant à notre travail pourrait être apporté par l'étude du compartiment spontané. En effet, l'étude de flux de gènes entre l'amandier cultivé et des populations de *P. orientalis* au Liban, Syrie et Turquie effectuée par Delplancke et al., (2012) a montré l'importance des flux de gènes inter spécifique. De plus, notre travail a montré l'importance de l'usage qu'en font les agriculteurs. Donc, un travail plus approfondi sur la diversité de *P. orientalis* serait intéressant de même qu'une évaluation de l'existence éventuelle d'hybridation avec *P. spartioides*.

iv) Utilisation durable

L'utilisation durable des ressources phylogénétiques à travers la sélection de clones élites et la mise en place de programme d'amélioration génétique des espèces cultivées constitue un objectif spécifique de la stratégie nationale (cf. encadré). Des cultivars d'amandier ont été sélectionnés (et par la suite conservés) à travers le temps par les agriculteurs. Dans une optique d'intégrer conservation et valorisation, la mise en place d'un programme de sélection participative est envisageable avec la catégorie des agriculteurs-pépiniéristes identifiés dans la plaine de la Bekaa et à Akkar (Liban Nord) pour développer une utilisation durable des ressources génétiques de l'amandier. Un travail complémentaire pourrait aussi être engagé pour la prise en compte du compartiment spontané ; en particulier l'utilisation de *P. orientalis* comme porte-greffe.

Bibliographie

- Brush, S. B. (1991) 'A Farmer-based approach to conserving crop germplasm', *Economic Botany*, 45(2), pp. 153–165. doi: 10.1007/BF02862044.
- Chalak, L. (2015) 'National Strategy for Conservation and Management of Plant Genetic Resources for Food and Agriculture in Lebanon 2015-2035', *Food and Agriculture Organization of the United Nations*, (TCP/SNO/3401).
- Grasselly, C., Dirlewanger, E. (2016) '1978 / Lauranne : la première variété autofertile d'amandier commercialisée', *INRA Science & Impact*. Available at: <http://www.bordeaux->

aquitaine.inra.fr/Toutes-les-actualites/70-ans-Lauranne-premiere-variete-amandier-autofertile-commercialisee (Accessed: 11 July 2018).

Choueiri, E., Verdin, E. Danet, J. L., Jreijiri, F., El Zammar, S., Salard, P., Bové, J.M., Garnier, M. (2003) *A phytoplasma disease of almond in Lebanon [2003]*, *AGRIS: International Information System for the Agricultural Science and Technology*. Available at: <http://agris.fao.org/agris-search/search.do?recordID=QC2003001784> (Accessed: 11 July 2018).

Delplancke, M. *et al.* (2012) 'Gene flow among wild and domesticated almond species: Insights from chloroplast and nuclear markers', *Evolutionary Applications*, 5(4), pp. 317–329. doi: 10.1111/j.1752-4571.2011.00223.x.

Delplancke, M. *et al.* (2013) 'Evolutionary history of almond tree domestication in the Mediterranean basin', *Molecular Ecology*. Wiley/Blackwell (10.1111), 22(4), pp. 1092–1104. doi: 10.1111/mec.12129.

Gouraud, Services du général (1921) 'La culture des arbres fruitiers en Syrie et Cilicie', *Revue de botanique appliquée et d'agriculture coloniale*, 1, pp. 129–148. doi: <https://doi.org/10.3406/jatba.1921.4001>.

Grasselly, C. and Crossa-Raynaud, P. (1980) *L'amandier*. Maisonneuv.

Hafizi, A. *et al.* (2013) 'Identification of new S-RNase self-incompatibility alleles and characterization of natural mutations in Iranian almond cultivars', *Trees*. Springer-Verlag, 27(3), pp. 497–510. doi: 10.1007/s00468-012-0803-7.

Haouane, H. (2012) *Origines , domestication et diversification variétale chez l ' olivier (Olea europaea L .) à l ' ouest de la Méditerranée*.

Heckathorn, D. D. (1997) 'Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations*', *Source: Social Problems*, 44(2), pp. 174–199. Available at: <http://www.jstor.org/stable/3096941> (Accessed: 11 July 2018).

Jarvis, D. and Hodgkin, T. (2000) 'Farmer decision making and genetic diversity : linking multidisciplinary research to implementation on-farm', in *Brush, S.B. (Ed.) Genes in the Field: On-Farm Conservation of Crop Diversity*. Lewis Publishers, Boca Raton, FL, USA. IDRC, Ottawa, ON, CA, pp. 261–278. Available at: <https://idl-bnc-idrc.dspacedirect.org/handle/10625/30843> (Accessed: 11 July 2018).

Kodad, O. *et al.* (2010) 'Molecular characterization of five new S alleles associated with self-incompatibility in local Spanish almond cultivars', *Options Méditerranéennes*, 94, pp. 105–109. Available at: https://citarea.cita-aragon.es/citarea/bitstream/10532/1762/1/2011_080.pdf (Accessed: 11 July 2018).

Marchenay, P. (2005) 'Conserver vivant, savoirs et pratiques locales : une gageure ? Le défi de la simultanéité', *Biodiversité et savoirs naturalistes locaux en France Laurence*, in Bérard L. Cegarra M., Djama M., Louafi S., Marchenay P. Roussel B. Verdeaux F (eds), CIRAD - IDDRi - IFB- INRA: 91–98. <http://www.ethno-terroirs.cnrs.fr/IMG/pdf/tapMarchenay.pdf>

Annexes

Annexe 1

Enquête agriculteurAmandier

Date :

Sites (Régions) :

Confession :

Nom et prénom du propriétaire :

N° tel :

Nom et prénom de la personne entretenue :

N° tel :

Description du verger :

- | | | |
|----------------------------|-----------|-------------------------------------|
| • Pente : | | Age : |
| • Négligé | Abandonné | Entretenue (façon d'irrigation) : |
| • Cultivé : | | Sauvage : |
| • Semis (type du graine) : | | Greffage (sur quel porte greffe) : |

Planté par :Activité dans le site (bergers, propriétaire, touristes, présence de ruches) :Origine des plantes

Importée (pays d'origine) :

Locale (région d'origine) :

Commercialisation :

Exportation, pour quelles pays :

Marché locale, pour quelles régions :

Récolte :

- | | | |
|---------------------|------------------|----------------------|
| • Unique : | | A différents temps : |
| • Vert | Semi-sec (ferki) | Sec |
| • Date du récolte : | | |

Consommation :

- | | | |
|--------|------------------|-----|
| • Eté | Hiver | |
| • Vert | Semi-sec (ferki) | Sec |

Usage, Recettes :

Echange du matériel Végétale :

OUI

NON

- Type d'échange (entre familles, amis...) :
- Echange au sein du village :
Echange entre villages proches:
Pas de limites d'échange :

Connaissances :

1- En fonction de quoi choisissez-vous ?

- La graine du semis

Type choisi :

Cultivée, douce

Spontanée, amer

- Le porte greffe

Type choisi :

Cultivée, douce

Spontanée, amer

- Le greffon

Type choisi:

2. Pour quel raison avez-vous choisi les variétés plantés dans votre terrain ?

3. Est-ce qu'il y a des limites de plantation pour l'amandier ? limites de transport, en distance ou en hauteur, du greffon ou du porte greffe ?

a. Est-ce qu'on greffe sur les arbres spontanés âgés ?

4. Age limite de l'amandier dans la région ?

5. Attaque du gel ?

6. Maladies présentes ?

7. Remplacement par une autre culture ? Laquelle et pour quel raison ?

8. Est-ce que vous avez des livres, des souvenirs ou des histoires concernant les amandiers ?

9. Est-ce que l'amandier fait partie de certaines traditions (confessionnelle ou/et sociale) ? Si la vie sociale ou la confession a influencée la culture des amandiers en générale et le choix variétale en particulier.

10. Est-ce-qu'il y a de ruches d'abeilles dans la région ?

11. Variétés connus selon les dires des paysans :

12. Autres :

Annexe 2

أستمارة مشتل

Enquête pépiniériste / amandier

- ✓ **Personne** / رقم الشخص: () PE
- ✓ **L'une des 2 Variétés cultivées O/N** / هل يوجد لديكم الصنف "حاواني", "خشابي"
- ✓ **Localite** / مكان وجود المشتل
- ✓ **Date de creation pepiniere** / سنة الانشاء
- ✓ **Age du pepinieriste** / عمر صاحب المشتل
- ✓ **Motivation à cette creation** / سبب التأسيس
- ✓ **Principales plantes multiplies** / ابرز النباتات الموجود
- ✓ **Part arbres fruitiers** / نسبة الأشجار المثمرة
- ✓ **Part amandiers** / نسبة اللوز
- ✓ **Intérêt d'une pépinière amandier** / أهمية مشتل اللوز
- ✓ **Principaux clients achetant des amandiers (et lieux)** / أبرز الزبائن التي تشتري منكم شتول اللوز
- ✓ **Origines des variétés d'amandier multipliées (lieux, personnes)** / مصدر أصناف اللوز المتكاثرة في المشتل
- ✓ **Modes d'acquisition des variétés d'amandier (achats, autres)** / طريقة الحصول على هذه الأصناف
- ✓ **Pratiques de propagation/multiplication des variétés d'amandier** / تقنية التكاثر المستخدمة في انتاج أصناف اللوز
- ✓ **Pratique de greffage des variétés d'amandier** / كيفية القيام بتقنية التطعيم ولأي أصنف
- ✓ **Contact autres pépiniéristes (O/N) : lieux, nom.** / هل تعاملون مع مشاتل أخرى: المكان و اسم المشتل

أستمارة مزارع

Réseau de diffusion de l'amandier cultivé

- ✓ **Variétés multipliée** / أصناف اللوز المتكاثرة في المشتل
- ✓ **Origine** / مصدر كل صنف: (Don , Achat) هبة أو شراء
- ✓ **Acquisition** / شكل التملك: (Plant, Graine) غرسة أو بذرة
- ✓ **Origine géographique** / المصدر الجغرافي:
- ✓ **Nom ou statut du "donneur" ou du vendeur** / اسم أو حالة الواهب أو البائع:
- ✓ **Prix payé OU contrepartie** / القيمة المدفوعة أو مقابل مادي أو نباتي أو معنوي أو هدية بلا أي مقابل:
- ✓ **Motivation de l'acquisition** / حافظ اقتناء الصنف: (compenser dégats et risques/diversifier le verger pour raisons économiques/croiser pour accroître rentabilité/ en vue d'un trait agronomique ou morphologique particulier)
- ✓ **Perspectives de gestion (meilleure production...)** رسم منظوري لإدارة مستدامة بهدف الحفاظ على التنوع

البيولوجي.

Résumé

Cette thèse est une contribution à la compréhension de l'influence de l'Homme sur la dynamique de la diversité génétique de l'amandier *Prunus dulcis* au Liban, centre de diversification de l'espèce. Cette espèce auto-incompatible a longtemps été propagée par semis, puis par greffage. La démarche interdisciplinaire suivie fait appel à la biologie des populations et à l'anthropologie.

Cette étude comprend trois parties: dans la première partie, nous avons cherché à comprendre les processus de diversification de l'amandier cultivé en se basant sur des évidences archéologiques puis historiques et sur les connaissances sur la biologie évolutive de l'amandier dans le Bassin Méditerranéen. Nos résultats permettent de distinguer l'impact de différentes périodes d'introduction d'amandiers au Liban sur la structure de la diversité génétique des vergers actuels.

La deuxième partie, a permis d'évaluer l'importance culturelle de l'amandier cultivé au Liban, et d'identifier sa diversité intra-spécifique telle qu'elle est perçue par les informateurs. Nos résultats montrent une hétérogénéité des connaissances des informateurs qui nous a permis de mettre en évidence une taxonomie locale flexible. Nous avons identifié deux variétés phares.

Dans la troisième partie, nous avons évalué l'effet du changement de pratiques de propagation sur la structuration et la dynamique de la diversité génétique. Nos résultats montrent que la diversité génétique du cultivar traditionnel « Khachabi », propagé par semis, est structurée géographiquement. Par contre, la diversité du second cultivar, « Halwani » introduit anciennement, est structurée en deux pools génétiques en lien avec les pratiques mixtes de multiplication qui incluent semis et greffage.

Cette thèse montre comment l'évolution des pratiques interagit avec la biologie de l'amandier pour en façonner la diversité. L'amandier a longtemps été propagé par voie sexuée. L'introduction du mode de propagation par greffage a été adoptée graduellement. Les agriculteurs ont ainsi maintenu une certaine diversité génétique par la multiplication sexuée occasionnelle des cultivars introduits. Les introductions tardives après l'adoption du greffage ont réduit la diversité. Nous illustrons ainsi comment les connaissances, les pratiques et les prises de décisions de l'Homme à différents échelles spatiales et temporelles influence la structure et la dynamique de la diversité des plantes cultivées.

Mots clés: Amandier, Diversité variétale, Pratiques culturelles, Liban.

Summary

This thesis is a contribution to the understanding of the influence of Human on the dynamics of the genetic diversity of the almond tree *Prunus dulcis* in Lebanon, center of diversification of the species. This self-incompatible species has long been propagated by seed and then through grafting. The interdisciplinary approach followed uses population biology and anthropology. This study consists of three parts:

In the first part, we sought to understand the processes of diversification of the cultivated almond tree based on archaeological and historical evidences and on the evolutionary biology of the almond tree in the Mediterranean Basin. Our results make it possible to distinguish the impact of the different periods of introduction of almond tree in Lebanon on the structure of the genetic diversity of young orchards.

The second part assessed the cultural importance of the almond tree grown in Lebanon, and identified its intraspecific diversity as perceived by the informants. Our results show heterogeneity of informants' knowledge that allowed us to highlight a flexible local taxonomy. We have identified two flagship varieties.

In the third part, we assessed the effect of the change in propagation practices on the structuring and dynamics of genetic diversity for these flagship varieties. Our results show that the genetic diversity of the traditional cultivar "Khachabi", propagated by seed, is geographically structured. On the other hand, the diversity of the second cultivar, "Halwani" introduced recently, is structured in two genetic pools in connection with the mixed practices of multiplication which includes sowing and grafting.

This thesis shows how the evolution of practices interacts with the biology of the almond tree to shape its diversity. The almond tree has long been propagated sexually. The introduction of the graft propagation mode has been adopted gradually. Farmers thus maintained a certain genetic diversity through the occasional sexual multiplication of introduced cultivars. Late introductions, after grafting adoption, have reduced diversity in recent orchards.

We illustrate how human knowledge, practices and decision-making at different spatial and temporal scales influence the structure and dynamics of plant diversity.

Keywords: Almond, varietal diversity, cultural practices, Lebanon.